

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2005, 03:00:42 ; Search time 31 Seconds  
(with about 31 iterations)

3236.398 million cell updates/sec

Title: US-10-066-521-6

Sequence: 1 MEGDKSLTFSSYGLQWCLYE.....DDHSGVSWSLGAGLEGLVS 1344

Scoring table: BLOSUM62

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

```
Minimum DB seq length: 0
```

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

**Listing first 45 summaries**

Database : Issued Patents AA: \*

- 1 /cgn2\_6/ptodataa/1/iaa/5A COMB. pep: \*
- 2 /cgn2\_6/ptodataa/1/iaa/5B COMB. pep: \*
- 3 /cgn2\_6/ptodataa/1/iaa/6A COMB. pep: \*
- 4 /cgn2\_6/ptodataa/1/iaa/6B COMB. pep: \*
- 5 /cgn2\_6/ptodataa/1/iaa/PCTUS COMB. pep: \*
- 6 /cgn2\_6/ptodataa/1/iaa/backtitled. pep: \*

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

No.	Score	Query	Length	DB	ID	Description
1	945.5	13.4	1454	4	US-09-368-221B-10	Sequence 10, Appl1
2	945.5	13.4	1473	4	US-09-368-221B-2	Sequence 2, Appl1
3	938.5	13.3	1399	4	US-09-368-221B-4	Sequence 4, Appl1
4	938.5	13.3	1424	4	US-09-368-221B-12	Sequence 12, Appl1
5	938.5	13.3	1443	4	US-09-368-221B-6	Sequence 6, Appl1
6	630.5	8.9	4456	2	US-08-910-731-4	Sequence 4, Appl1
7	630.5	8.9	4556	2	US-08-795-395-4	Sequence 4, Appl1
8	621	8.8	4556	2	US-08-910-731-2	Sequence 2, Appl1
9	621	8.8	4562	2	US-08-795-395-2	Sequence 2, Appl1
10	612	8.7	4562	2	US-08-910-731-8	Sequence 8, Appl1
11	607.5	8.6	4612	2	US-08-910-731-6	Sequence 6, Appl1
12	374.5	5.3	953	3	US-09-245-281-43	Sequence 43, Appl1
13	374.5	5.3	953	4	US-09-207-359B-43	Sequence 43, Appl1
14	374.5	5.3	953	4	US-09-340-620A-43	Sequence 43, Appl1
15	374.5	5.3	953	4	US-09-865-364-43	Sequence 43, Appl1
16	364	5.1	1013	4	US-10-014-269-3	Sequence 3, Appl1
17	364	5.1	1040	4	US-10-014-269-2	Sequence 2, Appl1
18	364	5.1	1043	4	US-09-949-016-9080	Sequence 9080, Appl1
19	363	5.1	1040	4	US-10-014-269-4	Sequence 4, Appl1
20	356.5	5.0	209	4	US-09-368-221B-19	Sequence 19, Appl1
21	347	4.9	953	3	US-09-029-041A-8	Sequence 8, Appl1
22	347	4.9	953	3	US-09-245-281-8	Sequence 8, Appl1
23	347	4.9	953	4	US-09-207-359B-8	Sequence 8, Appl1
24	347	4.9	953	4	US-09-340-620A-8	Sequence 8, Appl1
25	347	4.9	953	4	US-09-865-364-8	Sequence 8, Appl1
26	345	4.9	1007	4	US-10-014-269-34	Sequence 34, Appl1
27	324.5	4.6	966	4	US-09-207-359B-47	Sequence 47, Appl1

28	324.5	4.6	96.6	4	US-09-865-344-47	Sequence 17, Appl
29	306.5	4.3	83.0	4	US-09-064-199-19	Sequence 19, Appl
30	306.5	4.3	110.6	4	US-09-064-199-17	Sequence 17, Appl
31	306.5	4.3	113.0	2	US-08-513-547A-6	Sequence 6, Appl
32	306.5	4.3	113.0	4	US-09-064-199-18	Sequence 18, Appl
33	306.5	4.3	120.7	4	US-09-064-199-16	Sequence 16, Appl
34	213	3.0	48.3	4	US-09-904-615-154	Sequence 15, Appl
35	163	2.3	30.5	4	US-10-014-269-7	Sequence 7, Appl
36	160	2.3	124.0	3	US-08-930-996A-4	Sequence 4, Appl
37	153.5	2.2	40.0	3	US-09-099-041A-26	Sequence 26, Appl
38	153.5	2.2	45.0	3	US-09-245-281-26	Sequence 26, Appl
39	153.5	2.2	40.0	4	US-09-207-359B-26	Sequence 26, Appl
40	153.5	2.2	40.0	4	US-09-340-620A-26	Sequence 26, Appl
41	153.5	2.2	40.0	4	US-09-865-364-26	Sequence 26, Appl
42	149	2.1	50.9	3	US-08-809-999D-17	Sequence 17, Appl
43	149	2.1	50.9	3	US-09-065-437-17	Sequence 17, Appl
44	149	2.1	50.9	3	US-09-322-360-17	Sequence 17, Appl
45	149	2.1	50.9	3	US-09-131-831B-17	Sequence 17, Appl
GENERAL INFORMATION						

## ALIGNMENTS

```

RESULT 1
US-09-388-221B-10
; Sequence 10, Application US/09388221B
; Patent No. 6618750
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221B
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1454
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Construct
US-09-388-221B-10

```

Query Match	13.4%;	Score 945.5;	DB 4;	Length 1454;
Best Local Similarity	24.8%;	Pred. No. 5.7e-83;		
Matches 294;	Conservative 162;	Mismatches 339;	Indels 389;	Gaps 30

```

0Y 11 SYGJWCIVE-LDKXEFQFKEKLLKKKS-SBSTSISIQFELNNANNECALLLHEXYGA 68
Db 5 AMERLACYLFEELKEELKEFOLLNKKASHRSSSETPAOPEKTSIGMEVASYLVAQ-YGE 63
0Y 69 SLAMATISISIFEMONJRTLSSEKARD----- 94
Db 64 QRAMDALHTWEMQJSLCAOAGAGHSBSPFYSPSEPHLGSBQPTSTAVLPMWIE 122
0Y 95 -----MKKISQAM----- 102
Db 124 LPACTGGSERRVLRDLPDTSGRWRREISASILLYALPSSFDHSPSOESNAPTSTAVL 182
0Y 103 -----BOEGA-----TAAETEOEISQAMEQBAGATAETEOQ 134
Db 184 GSWGSPRQPSLARBEQBARCTQMPRLDETSIGIYUETIRERERKEKBEKRRPMAAVGTPRQ 243
0Y 135 GHGG-----DTWYDKSHVMT-KF-----AEEEDVARSFENT 164
Db 244 AHTSLQPHNHMRWVSRYBSLSTGWKMDKEDFNQKCTQOLLQRRHPRNSQDPLVAKS----- 299
0Y 165 AADWP-----EMQTLAG-AFDSDRMGFRPTVVLHGKSGIGKSLARLIVLCMA 212
Db 300 ---WPUVEENRGHLLIEIRDLFGPGLDRO---EPRVILLAGAAGIGKSTLARQVKEAWG 352
0Y 213 QGGLYQGMSTSYVFLFVREMOJKKSSVTEFLSRKMPDSQAPVTEIMSRPRLLFTIDGF 272

```

Db 353 RGLYGRDFQHFVYFSCRELAQSKVSVSLAELIGKDTATPAPIRQILSRPERLLFILDGV 412  
Qy 273 DDLSGVANN-DTKCKDMAEKQPPFTLIRSLIRKTLBPESFLIYVVRVGTREKLSEVVS 331  
Db 413 DEPGVNLDEPSSSELCLHWSQOPADALLGSLGKLTLEBPASFLIYATTAQNLIPSEIQ 472  
Qy 332 PRYLVRIGISGEQRTHLLERIGIGEHOKTOGLRAIMNRRELLDQCOVPAGSLICVALQL 391  
Db 473 ARWVAVLGFSSSRKEFYRYFTDERQAIRAFRLVSKNKMALCLVWVSWLACTCIMO 532  
Qy 392 QDVVGSVAAPNQTLLG--LHAAFVHQLTTPRGVVRCLINEERVYLKRCMAVEGVN 449  
Db 533 QMKREKLTLSKTTTLCIHYLAQALQAPLG-----POLRDLCSLAEBGIWQ 581  
Qy 450 RKSVPDGDLMVQGLGSESELALFPHNLT-----LPDSHCEYTFPHLSLDDFCALYYV 505  
Db 562 KKTLPSPDLRKHGIDGAIISTFLMKGILQEHPIPLS-----YSTHLCFOQFFAAMSIV 636  
Qy 506 LEGLE--IEPALCPLYVEKTRSMELKQAGFHIHSLM--MKRFLGLVSEDPVRRPLEVL 560  
Db 637 LEDEKGRGKHSNCIIDEKTLLEA-----YGIHGLFGASTTRFLGLLSDGEREMENI 689  
Qy 561 LGCPVPLGVKQKLIHWVSLG--QOPNATTPGDTLDAPHCLFEIODEKFFVRLALNSPOEV 618  
Db 690 FHCRLSQG--RNLQMWVPSLQLLQPH-----SLESJHCLYETENKFTLQVMAHPEEM 741  
Qy 619 WLPINQNLDIASSFCLOHCPYLKIRVDVKGIFPDESABCPVPLVMDKTLIEBQW 678  
Db 742 GWCVTDMELVCTFCIKFSRHVKQLQ--IEG--RQHRSTWSPMTVVLFKRWVPTDAYW 797  
Qy 679 EDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLHRPTCKIOTLMFRNAQITPGVQHLW 738  
Db 798 -----QLTF 801  
Qy 739 RIWMANRNLASLNGTHLKEEDVYMACBALKHPCLLESRLDCCGLTHACYIKISQIL 798  
Db 802 SVLKATRNMLKEILDLSGNSLSHSAVSKLCTKLRRPRCLLETLR----- 843  
Qy 799 TTSPLKSLSLAGNKVTDQGVWPLSDALRVSCALQKLIIDECGITTATGCSLASALVSN 858  
Db 844 -----LAG-----CGITLADCKDLAFGLRAN 864  
Qy 859 RSLTHCLSNNSLGNVNLCSRMRLLPHCSLQRLMLNQCILDTPAGCGFLALALMGNSWL 918  
Db 865 QT-----L 867  
Qy 919 THLSLSTMPVEDNGYKLICEVWRBPSCHLODELVKCHLTAAACESLSCVISRSHKSL 978  
Db 868 TELDLSFVNLTDGAKHLCORLRQPSCKLQRLQVLSGCLTSDCCDLASVLSASPSLKEI 927  
Qy 979 DLTDLALDGGVVALCEGLKQKNSVLTGLGKACGLTSDCCBAL 1022  
Db 928 DLQONNLDDVGVRLCEGLRHPACKLIRLGLDQTTLSDEMROEL 971

RESULT 2  
US-09-388-221B-2  
; Sequence 2, Application US/09388221B  
; Patent No. 6818750  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; TITLE OF INVENTION: No. 6818750el Card Proteins Involved in Cell Death Regulation  
; FILE REFERENCE: P-LJ 3650  
; CURRENT APPLICATION NUMBER: US/09/388,221B  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1473  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-388-221B-2

Query Match 13.4%; Score 945.5; DB 4; Length 1473;  
Best Local Similarity 24.8%; Pred. No. 5.8e-83;  
Matches 294; Conservative 162; Mismatches 339; Indels 389; Gaps 30;  
Qy 11 SYGLQWCLYE-LDKEPOTFERELIKKKS-SESTTCSIPQEIENANVECLALLHEYYGA 68  
Db 5 AMGLIACYLEFKKEBEKEFOLLIANKAHSSSGETPAOPERTSGMEVASYLVAQ-YGE 63  
Qy 69 SLAATSTISIFENNLTLSKARD----- 94  
Db 64 QRADWLALHTWQGLKNSLCAQAOEGAGHSBPSPSPSEPHLSPSPOTSTAVLMPIWE 123  
Qy 95 -----MKKISQAM----- 102  
Db 124 LPACQTGSSERRVLRQLPDTSGRRWRIRASLVLQALPSPDHSPQESPNAPTSTAVL 183  
Qy 103 -----EQGCA-----TAETEROEISQAMEQCATAELEBQ 134  
Db 184 GSWGSPPOPSLAPREOAPGTQWPLDETSGIYYEIREREREKSGRPPMAAVVGTTPQ 243  
Qy 135 GHGG-----DTWDYKSHWT-KF-----AEEEDVRSFENT 164  
Db 244 AHTSLQPHHHPWSPVSRESLSTPWXEDPNQKFTQULLQRPHPQSODPLVKS----- 299  
Qy 165 AADWP-----EMQTLAG-APDSRWGFRPRTVVLHGKSGIKSALARRIVLCWA 212  
Db 300 ---WPDYVEENRGLHLEIRDLFGGLDTQ---EPRIVILQAGAGIKSTLARGVKEAMG 352  
Qy 213 QGGIYQGMFVSFVFLPYREMKRKSSVTEFISHEWDSQAPVTEINSRBRLLFITDGF 272  
Db 353 RGLYGRDFQHFVYFSCRELAQSKVSVSLAELIGKDTATPAPIRQILSRPERLLFILDGV 412  
Qy 273 DDLSGVANN-DTKCKDMAEKQPPFTLIRSLIRKTLBPESFLIYVVRVGTREKLSEVVS 331  
Db 413 DEPGVNLDEPSSSELCLHWSQOPADALLGSLGKLTLEBPASFLIYATTAQNLIPSEIQ 472  
Qy 332 PRYLVRIGISGEQRTHLLERIGIGEHOKTOGLRAIMNRRELLDQCOVPAGSLICVALQL 391  
Db 473 ARWVAVLGFSSSRKEFYRYFTDERQAIRAFRLVSKNKMALCLVWVSWLACTCIMO 532  
Qy 392 QDVVGSVAAPNQTLLG--LHAAFVHQLTTPRGVVRCLINEERVYLKRCMAVEGVN 449  
Db 533 QMKREKLTLSKTTTLCIHYLAQALQAPLG-----POLRDLCSLAEBGIWQ 581  
Qy 450 RKSVPDGDLMVQGLGSESELALFPHNLT-----LPDSHCEYTFPHLSLDDFCALYYV 505  
Db 562 KKTLPSPDLRKHGIDGAIISTFLMKGILQEHPIPLS-----YSTHLCFOQFFAAMSIV 636  
Qy 506 LEGLE--IEPALCPLYVEKTRSMELKQAGFHIHSLM--MKRFLGLVSEDPVRRPLEVL 560  
Db 637 LEDEKGRGKHSNCIIDEKTLLEA-----YGIHGLFGASTTRFLGLLSDGEREMENI 689  
Qy 561 LGCPVPLGVKQKLIHWVSLG--QOPNATTPGDTLDAPHCLFEIODEKFFVRLALNSPOEV 618  
Db 690 FHCRLSQG--RNLQMWVPSLQLLQPH-----SLESJHCLYETENKFTLQVMAHPEEM 741  
Qy 619 WLPINQNLDIASSFCLOHCPYLKIRVDVKGIFPDESABCPVPLVMDKTLIEBQW 678  
Db 742 GWCVTDMELVCTFCIKFSRHVKQLQ--IEG--RQHRSTWSPMTVVLFKRWVPTDAYW 797  
Qy 679 EDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLHRPTCKIOTLMFRNAQITPGVQHLW 738  
Db 798 -----QLTF 801  
Qy 739 RIWMANRNLASLNGTHLKEEDVYMACBALKHPCLLESRLDCCGLTHACYIKISQIL 798  
Db 802 SVLKATRNMLKEILDLSGNSLSHSAVSKLCTKLRRPRCLLETLR----- 843  
Qy 799 TTSPLKSLSLAGNKVTDQGVWPLSDALRVSCALQKLIIDECGITTATGCSLASALVSN 858  
Db 844 -----LAG-----CGITLADCKDLAFGLRAN 864  
Qy 859 RSLTHCLSNNSLGNVNLCSRMRLLPHCSLQRLMLNQCILDTPAGCGFLALALMGNSWL 918

```

Db 865 QT-----L 867
Qy 919 THLSLNNPVEDNGVKLLCEVWRBPSCHLODELVKCHLTAACCSLSCVISRSRHLSL 978
Db 868 TELDSFVNLTDAGAKHLCQRLRQPSCKLQRLQVSCGLTSDCCODLASVLSASPSLXEL 927
Qy 979 DLTNALGDGVAALCEGLKOKNSVLTRLGKACGLTSDCCAL 1022
Db 928 DLQNNLDDVGVRLLCEGLRHPRACKLIRLGLDQTTLSDEKQEL 971

RESULT 3
US-09-388-221B-4
; Sequence 4, Application US/09388221B
; Patent No. 6818750
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-IJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221B
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1399
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-388-221B-4

Query Match 13.3%; Score 938.5; DB 4; Length 1399;
Best Local Similarity 24.9%; Pred. No. 2,66-82;
Matches 291; Conservative 159; Mismatches 331; Indels 389; Gaps 30;

Qy 11 SYGLQWCLYE-LDKKEPOTFKELKKKS-SESTTCSIPQFIENANVECLALLHEYYGA 68
Db 5 AMGRACLAEFLFKELKEFOLLANKAHSRSSGETPAQPEKTSGMVASYLVAQ-YGE 63
Qy 69 SLAWATSISENNNLRTLSEKARD----- 94
Db 64 QRANDLALHTWQGLSLCAQAGAGHSPSPSPHILGSPSPSTAVLMPWIHE 123
Qy 95 -----MKKISQAM----- 102
Db 124 LPAGCTQSSERRVLAQLPDTSGRWRKEISALLYALPSSPDHSPSGSPNAPPTSTAVL 183
Qy 103 -----EOEGA-----TAETEOEISQAMEOEGATAETEOQ 134
Db 184 GSWSPSPQSLAPRQEAAPGTQWPLDSTSGIYTTIRERERKSKGRPPMAAVGTTPQ 243
Qy 135 GHGG-----DTWDYKSHVMT-KF-----AEBEDVRRSFENT 164
Db 244 AHTSLQPHHWPESVRESLCTWPMKNEDFNQFTQLLLQRPHPRSQDPLVKRS----- 299
Qy 165 AADWP-----EMQTLAG-AFDDRMGFRPRVTVLHKGSGIGSALARRIVLCMA 212
Db 300 ---WPDYVENRKHILIRIDLFEGPLDQ---EPRIVILQCAAGIGSTLARRQYKEMG 352
Qy 213 QGGLQGFVSYYFLPVRMOKKSSVTEFISRWPMQSAVPTIMSRPERLLFIIDGF 272
Db 353 RGQGLYGDFFQVHYFVSCRELAQSKVSLABELGKQCTATPARIIRILSRPERLLFIIDGV 412
Qy 273 DDLGSLVANN-DTKLCKDMAEKQPPFTLIRSLIRKVLBPESFLIVTRDVGTEKSEVVS 331
Db 413 DEPGVNLDEPSELCTHWSQPOPADALIGSLIGKITLPEASPLIATARTALQNLIPSEIQ 472
Qy 332 PRYLLVRISGEQRTHLLERGIGEHOKTQGLRAIMNRRELLDQCOVAVAGSLICVALQL 391
Db 473 ARWEVILPSSSSREKRYRYFTDRQALRAFRLVKSNEKELMALCLVWVWVSLACTCLMQ 532
Qy 392 QDVVGESEVAPFNQTLTG--LHAFAFVHQLTPRGVVRCLNLEERVLKRFQMAVEGVNN 449
Db 533 QMKREKXLTLSKTTTCTTCLHLYLAQALQAPLG-----POLMDLCSLAAGTIQ 581

```

```

Qy 450 RKSVPDGDLLWVQGLGSESELALFHMNL-----LPDSHCBEYTYTFHLSLDQFCALYYV 505
Db 582 KTLFSPDDLKRGHGDGAIISTFLAKGILQHPILS-----YSTHLCFQEFPAANSYV 636
Qy 506 LEGLE--TEPALCPLYEKTYSRMELKQGFHISLW---MKRFLFGLVSDVRPRLEVL 560
Db 637 LEDEGRKRNKNCIIDEKTELEA-----YGHGLFQASTTRFLGLISDEGREMENI 689
Qy 561 LCGPVPLGVKQKLLHWVSLG--QOPNATTPGDTLDAFHLCPETQDKEFVRLALNSFOEV 618
Db 690 FHCRLSQG--RNLQWVPSLQTLQPH-----SLESILHCLYETRNKFTQVMAHFEEM 741
Qy 619 WLPIQNQNDLASSRCLQHCYLAIRKIVDVGIFPRDESACAPVPLVMDKTLIEQW 678
Db 742 GWCVTDEHLLVCTFCIKFSRHVKLQD-LIG---RQHRSTWSPMTVVLPFRVWPVTDAYW 797
Qy 679 EDFCSMLGTHPHRLQDLGSSILTERAMKTLCAKLHPTCKIQTLMPRNAQITPGVQHLW 738
Db 798 -----QLLF 801
Qy 739 RIVANRNRLSLNIGTHLKEEDVYMACALKHPKCLLESRLDCCGLTHACYLKSQIL 798
Db 802 SVLAKTRMLKELDPSGSLSHSAVKSICKTLRRPRCLLETIR----- 843
Qy 799 TTSPSLKSLAKGKRVTDQGMPLSDALRVSCALQKILBEDCGITATGQSLASALVSN 858
Db 844 -----LAG-----CGILMEDCKDLAFGLFAN 864
Qy 859 RSLTHLCLSNNSLQNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTPAGCGFLALALMGNWL 918
Db 865 QT-----L 867
Qy 919 THLSLNNPVEDNGVKLLCEVWRBPSCHLODELVKCHLTAACCSLSCVISRSRHLSL 978
Db 868 TELDSFVNLTDAGAKHLCQRLRQPSCKLQRLQVSCGLTSDCCODLASVLSASPSLXEL 927
Qy 979 DLTNALGDGVAALCEGLKOKNSVLTRLG 1008
Db 928 DLQNNLDDVGVRLLCEGLRHPRACKLIRLG 957

RESULT 4
US-09-388-221B-12
; Sequence 12, Application US/09388221B
; Patent No. 6818750
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-IJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221B
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 1424
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Construct
US-09-388-221B-12

Query Match 13.3%; Score 938.5; DB 4; Length 1424;
Best Local Similarity 24.9%; Pred. No. 2,76-82;
Matches 291; Conservative 159; Mismatches 331; Indels 389; Gaps 30;

Qy 11 SYGLQWCLYE-LDKKEPOTFKELKKKS-SESTTCSIPQFIENANVECLALLHEYYGA 68
Db 5 AMGRACLAEFLFKELKEFOLLANKAHSRSSGETPAQPEKTSGMVASYLVAQ-YGE 63
Qy 69 SLAWATSISENNNLRTLSEKARD----- 94
Db 64 QRANDLALHTWQGLSLCAQAGAGHSPSPSPHILGSPSPSTAVLMPWIHE 123

```

QY 95 -----MKKISQAM----- 102  
DB 124 LPAGCTGSSERRVLRQLPDTSGRRKREISASLIYQALPSSPDHSPSOSPNAPTSTAVL 183  
QY 103 -----EODEGA-----TAETEOEISQAMEOGATTAETEO 134  
DB 184 GWSGSPQPSLAPRQGAFTQWPLDETSGIYTEIRERREKSEKGRPPMAAVGTPTQ 243  
QY 135 GHGG-----DTWDYKSHWT-KF-----AEEEDVARSFENT 164  
DB 244 AHTSLQPHHNPWPSVRESLCSYTWPMKXEDFNQKFTQLLLQRPHPRSODPLVRS----- 299  
QY 165 AADWP-----EMQTLAG-AFSDRWGFRPRTVVLHGSGIGSKALARRIVLCMA 212  
DB 300 ---WPDYVENRGHILIEIRDLFGSLDTQ---EPRIYILQGAAGIGSTLARQVKEMWG 352  
QY 213 QGGIYQGMFSYVFLPVREMOQKKESSVTEFISREWPDSQAPVTEIMSPERLLFITIDG 272  
DB 353 RGQLYGDRFGHFYFSCRELAQSKVSLAEILGKDGATAPAPIRQILSRPERLLFITIDG 412  
QY 273 DDLSGLVANN-DTKLCKMAKQRPFTLIRSLRKVLLEPSFLIYTVRVGTREKLSSEVVS 331  
DB 413 DEPGWVLOEPSESELCHWSQOPADALGSLGKTLIEBASFLITARTTALQNTLIPSLQ 472  
QY 332 PRYLLVRIGISGEORIHLLBERGIGEHOKTOGLRAIMNNREILDOCVAVGSLICVALQL 391  
DB 473 ARWVEVLGFSESSKKEFYRYFTDERQALRAFLVYSKNEIMALCLVWVSWLACTCLMQ 532  
QY 392 QDVVGESVAPFNQTLTG--LHAAVFVHQLTTPRGVVRCLNBERVVLKRCMAVEGWN 449  
DB 533 QMKREKTLTTSKTTTTLCHYLAQALQAPLG-----POLADLCSLAABEGIMQ 581  
QY 450 RKSVDGDDLMVQIGSESELALFHMNTL---LPDSCEBYTTFPHLSIDPFCAALYV 505  
DB 582 KKTLPSPDDLRKGLDGAIIISTFLMKGILQEHPIPLS-----YSFHLCPQEFPAANSYV 636  
QY 506 LEGLE--IBPALCPLYVEKTKRSMELKQAGFHISLW---MKRFLFGLVSDVRPRPEVL 560  
DB 637 LEDEKGRKHSNCTIIDLEKTLFA-----YGHGLFGASTTFRLLGILSDEGEREMENI 689  
QY 561 LGCVPVLGVKOKLHMVSLLG--QQPNATTPGDTLDAFHCLFTQDKFVRLALNSFOEV 618  
DB 690 FHCRLSQC--RNLQMOWPSLQILLQPH-----SLESILCHLETYRNKFTLQVMAHFBEM 741  
QY 619 WLPINQNDLIASSPCLOHCEYLRKIRYDKGIFPRDSABACPVVPLMWDKTLIEBQW 678  
DB 742 GMCVETDMEILLVCTFCIKFSRHVKQLQI-IEG--RQHRSTWSPFMVVLFRWVPVTDAYW 797  
QY 679 EDFCSMLGTHPHLRQDLGSSILTERAMKCAKLRHPYCKIOTLMFRNAQITPGVOHLM 738  
DB 798 -----QTLF 801  
QY 739 RIVANRNRLRSNLGTHLKEEDVAMACEALKNPKCLLESRLDCCGILTHACYLKISQIL 798  
DB 802 SVLKATYRLKELDELSGNSLSHSAVSKLCTKLRRPRCLLETJLR----- 843  
QY 799 TTSFSLKSLAIGNKVTYDQGVMPISDALRVSCALQKILBEDCGITANGCSLASALVSN 858  
DB 844 -----LAG-----CGITADCDXDLAFGLRAN 864  
QY 859 RSLTHLCLSNNSLGNBGNLLCRSMRLPHSGLSQRLMNLQCHLDTAGCGFLALALMGNSWL 918  
DB 865 QT-----L 867  
QY 919 THLSLNNPVEDNGVYKLCCEVWRBPSCHLODELVYKCHLTAACESLSVLSRSRHLKSL 978  
DB 868 TELDLSFVVLTDAGAKHLCQRLRQPSCTLQRLQLYVSGLTSDCCQDLASVLSASPSLKEI 927  
QY 979 DLTNALADGGVAALCEGLKQKNSVLTPLG 1008  
DB 928 DLQGNLDDVGVRLCEGLRHPACKLIRLG 957

RESULT 5  
US-09-388-221B-6  
; Sequence 6, Application US/09388221B  
; Patent No. 6818750  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; TITLE OF INVENTION: No. 6818750el Card Proteins Involved in Cell Death Regulation  
; FILE REFERENCE: P-1J 3650  
; CURRENT APPLICATION NUMBER: US/09/388, 221B  
; NUMBER OF SEQ IDS NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1443  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-388-221B-6

Query Match 13.3%; Score 938.5; DB 4; Length 1443;  
Best Local Similarity 24.9%; Pred. No. 2,76-82;  
Matches 291; Conservative 159; Mismatches 331; Indels 389; Gaps 30;

QY 11 SYGLQWCLYE-LDKKEPQTFKELKKKS-SESTCSIPQEIENANVECLALLHEYYGA 68  
DB 5 AMGRACYLEFLKKEELKEFQILLANKAHSRSSGFTPADEKTSGMENVASYVAQ-YGE 63  
QY 69 SLAATGISIFENNTLTLSEKARD----- 94  
DB 64 QRADWLALHTWEQMLKSLCAQAOEGAGHSPSPSPSEPHLSPSQPTSTAVLMPWIEH 123  
QY 95 -----MKKISQAM----- 102  
DB 124 LPAGCTGSSERRVLRQLPDTSGRRKREISASLIYQALPSSPDHSPSOSPNAPTSTAVL 183  
QY 103 -----EODEGA-----TAETEOEISQAMEOGATTAETEO 134  
DB 184 GWSGSPQPSLAPRQGAFTQWPLDETSGIYTEIRERREKSEKGRPPMAAVGTPTQ 243  
QY 135 GHGG-----DTWDYKSHWT-KF-----AEEEDVARSFENT 164  
DB 244 AHTSLQPHHNPWPSVRESLCSYTWPMKXEDFNQKFTQLLLQRPHPRSODPLVRS----- 299  
QY 165 AADWP-----EMQTLAG-AFSDRWGFRPRTVVLHGSGIGSKALARRIVLCMA 212  
DB 473 ARWVEVLGFSESSKKEFYRYFTDERQALRAFLVYSKNEIMALCLVWVSWLACTCLMQ 532  
QY 392 QDVVGESVAPFNQTLTG--LHAAVFVHQLTTPRGVVRCLNBERVVLKRCMAVEGWN 449  
DB 533 QMKREKTLTTSKTTTTLCHYLAQALQAPLG-----POLADLCSLAABEGIMQ 581  
QY 450 RKSVDGDDLMVQIGSESELALFHMNTL---LPDSCEBYTTFPHLSIDPFCAALYV 505  
DB 582 KKTLPSPDDLRKGLDGAIIISTFLMKGILQEHPIPLS-----YSFHLCPQEFPAANSYV 636  
QY 506 LEGLE--IBPALCPLYVEKTKRSMELKQAGFHISLW---MKRFLFGLVSDVRPRPEVL 560  
DB 637 LEDEKGRKHSNCTIIDLEKTLFA-----YGHGLFGASTTFRLLGILSDEGEREMENI 689  
QY 561 LGCVPVLGVKOKLHMVSLLG--QQPNATTPGDTLDAFHCLFTQDKFVRLALNSFOEV 618



```

Db 690 FHCRISOG--RNIQMWPESLQILLQPH-----SLESILHLYETRNKFTLTQVMAHFEBM 741
Qy 619 WLPINQNDLTLASSFCLOHCPYLRKIRVDVKGIFPRDESAECPPVPLMRKTLIEBQM 678
Db 742 GNCVETDELLCTCFIKFSRHVKLQL-IEG---RQHRSTWSPTMVVLFRWVPVTDAYW 797
Qy 679 EDFCSMLGTHPHRLQDLGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQTTPGVQHLW 738
Db 798 -----QILF 801
Qy 739 RIWMANRLRSINLGTHLKEEDVMAECALHPKCLLESRLDCCGLTHACYLKSQIL 798
Db 802 SVLKATRYMLKELDLSGNSLSHSAVSKLCTKLRRPCLLETL----- 843
Qy 799 TTSPLKSLISLAGNVTDOGVWPLDALRVSCALQKILBCCGTTATGCGSLASALVSN 858
Db 844 -----LAG-----CGLTAECCOLAGLGRAN 864
Qy 859 RSLTHLCLSNNSLNEGAVLLCRSMRLPHCSLQRLMLNQCHLDTAGCGFLALAMGNSWL 918
Db 865 QT-----L 867
Qy 919 THLSLNMNPVEDNGVYLLCEVNRBPSCHLDLELYKCHLTAACCSLSCVYSRHLKSL 978
Db 868 TELDSFVYLTDAKANHLQRLRPSCKLQRLQVSCGLTSDCCODLASVLSASPSLKL 927
Qy 979 DLTNALGDGVYALCEGLKOKNSVYTRIG 1008
Db 928 DLQNNLDVGVYRLCEGLRHPACKLIRIG 957

```

RESULT 6  
US-08-910-731-4  
Sequence 4, Application US/08910731  
Patent No. 5932440  
GENERAL INFORMATION:  
APPLICANT: CHATTERJEE, DEB K.  
TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., N.W., SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/910,731  
FILING DATE: (Herewith)  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/795,395  
FILING DATE: 04-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/794,546  
FILING DATE: 03-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/024,057  
FILING DATE: 16-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: ESMOND, ROBERT W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0942.3440003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-910-731-4

```

Query Match 8.9%; Score 630.5; DB 2; Length 456;  
Best Local Similarity 34.0%; Pred. No. 8,5e-53;  
Matches 147; Conservative 75; Mismatches 201; Indels 9; Gaps 3;

```

Qy 659 EACPVPVPM-----MRDKTLIEBQWEDFCSMLGTHPHRLQDLGSSILTERAMKTLCAKLR 714
Db 17 ELPLPIQGYVVRDLDCCGLETRCDDISALQANPALTELRLNELDAGAGVVLQGLQ 76
Qy 715 HPTCKIQTLMFRNAQIT-PGVQHLWRIWMANRLRSINLGTHLKEEDVMAECALKHPK 773
Db 77 NPTCKIQTLSLQNSCLTEAGCGVLPDVLRSLSTLREHLNDPNPLGDDEGLKILCEGLRBPQ 136
Qy 774 CLLESLRLDCCGLTHACTYKTSQILTTSPSLKSLSLAENKVTDOGVWPLDALRVSCAL 833
Db 137 CRLEKQLQLEYCNLTATSCPEPLASVLRVPRDERELVLSNNDHEAGIHTLCOGLKDSACQL 196
Qy 834 OKLILEDGIGATGCGSLASALVSNRSLTHLCLSNNSLNEGAVLLCRSMRLPHCSLQRL 893
Db 197 ESKLIENGCTISANCKDLDVVASKASLQELDLGSKNLGNTGIALCAGLLPSCRKRTL 256
Qy 894 MLNQCHLDTAGCGFLALAMGNSWLTHLSLNMNPVEDNGVYLLCEVNRBPSCHLDLELY 953
Db 257 WLMDCVTAEGCKDLCVLRKAKQSLKELSLAGNELKDGQQLCGSLLEPGQLESIMVWK 316
Qy 954 KCHLTAAACCSLSCVYSRHLKSLDTNALGDGVYALCEGLKOKNSVYTRIGLAKCG 1013
Db 317 TCSLTAAACCPHRCVSLTRNSSLFELQWSSNPLGDSGVVELKALGYPTVIVRVMLGDOD 376
Qy 1014 LTSDCESLTLALSCNRLTSLNLYONNFPSPKMKLCSAPACPTSNQIIGL-----WKM 1069
Db 377 VTDSGCSLATVLANRSLRELDLSNMCMDGNGVYQLLESILKQPSCTIQLVLYDIYWTMD 436
Qy 1070 QYPVQIRKLEE 1081
Db 437 EVEDQLRLLEE 448

```

RESULT 7  
US-08-795-395-4  
Sequence 4, Application US/08795395  
Patent No. 5965399  
GENERAL INFORMATION:  
APPLICANT: CHATTERJEE, DEB K.  
TITLE OF INVENTION: Cloning and Expression of Rat Liver and  
Porkine Liver Ribonuclease Inhibitor  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., N.W., SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/795,395  
FILING DATE: 04-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/024,057  
FILING DATE: 16-AUG-1996

ATTORNEY/AGENT INFORMATION:  
NAME: ESMOND, ROBERT W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0942.3440002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 456 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-795-395-4

Query Match 8.9%; Score 630.5; DB 2; Length 456;  
Best Local Similarity 34.0%; Pred. No. 8.5e-53;  
Matches 147; Conservative 75; Mismatches 201; Indels 9; Gaps 3;  
QY 659 EACPVPPLM---MRDKTLEIOWEDFCSMGTGPHRLQDLGSSILTERAKTKLCAKLR 714  
DB 17 ELPLHIOQYVVRDLDGCLTEVRCKDIRSAIQANPALTELRLNELGDAVGVLQGLQ 76  
QY 715 HPCKIQITMFRNAQIT-PGVQHLMRIYMANRNIRSLNLGGTHLKEEDVMACEALKHPK 773  
DB 77 NPTCKIQKLSIQNSILTEAGCGVLPDVLRSLSTRLRLHLDNPNPLGDEGLKLCGLRDPQ 136  
QY 774 CLLSRLRDCGGLTHACVLTAKISQILTTSPSLKSLIAGNKVTDGQVMPISDALRVSOAL 833  
DB 137 CRLEKQLLEYCNLTATSCBPLSVLRKVPDFKELVLSNNDPHEAGIHVLCGLKDSACQL 196  
QY 834 QKLLDEGATATGCGSLASALVSNRSLTHLCLSNNSLGNVNLCSMKLPKCSIQRL 893  
DB 197 ESLKLENGITISANCKDLDVVAASKASIQELDLSNKLKNTGIALCSGILLPCCRRLTL 256  
QY 894 MLNQCILDTCAGCGFLALALMGNSWLTSLISMNPVEDNGVKLCEVMEPSCHLQDELV 953  
DB 257 WLMQCDVYAGCKDLCRVLRKQSLKELSLAGNELKDEGAQLCESLLEPCQLESIMVK 316  
QY 954 KCHLTAAACCSLSCVYSRHLKSLDLTNALDGGVVALCEGLKQKNSVYTRGLAKAG 1013  
DB 317 TCSLTAAACPHFCSTVLTNSSLFELQWSSNPLGDSGVVLCALGYPTVLRVLMIGDCD 376  
QY 1014 LTPSCCEALSLALSCNRHLTSLNLYVONFSPKMMKLCSAFACPTSNQIIGL----WKW 1069  
DB 377 VTDSGCSLAVLANRSLRELDLSNNGMGNGVQLLESKQPSCTIQGLVLYDIWTD 436  
QY 1070 QYVQIRKLEE 1081  
DB 437 EVEDQLRALREE 448

RESULT 8  
US-08-910-731-2  
Sequence 2, Application US/08910731  
Patent No. 5932440  
GENERAL INFORMATION:  
APPLICANT: CHATTERJEE, DEB K.  
APPLICANT: SHANDILYA, HARINI  
TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., N.W., SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/910,731  
FILING DATE: (Herewith)  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/795,395  
FILING DATE: 04-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/794,546  
FILING DATE: 03-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/024,057  
FILING DATE: 16-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: ESMOND, ROBERT W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0942.3440003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 456 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-910-731-2

Query Match 8.8%; Score 621; DB 2; Length 456;  
Best Local Similarity 31.9%; Pred. No. 7.4e-52;  
Matches 151; Conservative 86; Mismatches 210; Indels 26; Gaps 5;  
QY 625 NLDIASSFCLOHCPYLRKIRVVDKGFPRDESACPVPLM---MRDKTLEIOWED 680  
DB 2 NLDI-----HCEQLSDAR-----WTELLPLIQYEVVRDLDGCLTEVRCKD 42  
QY 681 FCSMLGTHPHRLQDLGSSILTERAKTKLCAKLRHPCKIQITMFRNAQIT-PGVQHLMR 739  
DB 43 IGSALRANPSLTTELCTAFTELNELGDAVHVLVQGLQSPCKIQKLSIQNSILTEAGCGVLP 102  
QY 740 IWMANRLRLSLNLGGTHLKEEDVMACEALKHPKCLLESIRLDCCGTHACVLTAKISQILT 799  
DB 103 TLRSLPTLRHLSDNPLDAGRLCEGLDPCHEKQLLEYCRLTAAACCEPLASVLR 162  
QY 800 TSPSLKSLIAGNKVTDGQVMPISDALRVSOALQKLLDEGATATGCGSLASALVSNR 859  
DB 163 ATRALKELTVSNNDIGAGARVLGQGLADACQLETRLKNCGLTPANCODLSIVASQA 222  
QY 860 SLTHLCLSNNSLGNVNLCSMKLPKCSIQRLIMLNCCHLDTCAGCGFLALALMGNSWLT 919  
DB 223 SLRELDSNGLDGAGIARLCPGLSPASRLKTLMLWECIDITASGCRDLCRVLAQAKETLK 282  
QY 920 HLSLSMNPVEDNGVKLCEVMEPSCHLQDELVVKCHLTAAACCSLSCVYSRHLKSLD 979  
DB 283 ELSLAGNKLDDEGARLTCESLLOPGCQLESIMVKSCVLTAAACQHSVLMITONHLELQ 342  
QY 980 LITNALDGGVVALCEGLKQKNSVYTRGLKACGLTSDCCCEALSLALSCNRHLTSLNLYQ 1039  
DB 343 LSSNKLDSDGIQELCQALSQPGTTLRVLCIDGDEVNMGSSSLASILLANRSLRELDLSN 402  
QY 1040 NNFSPKMMKLCSAFACPTSNQIIGLMKWQYVQIRKLEEYQLKP--RVV 1090  
DB 403 NCVGDPVQLQLGSLBQPCALBQLVLYDTYMTVEVEDRLQALEGSRKGRVLI 455

RESULT 9  
US-08-795-395-2  
Sequence 2, Application US/08795395  
Patent No. 5965399  
GENERAL INFORMATION:  
APPLICANT: CHATTERJEE, DEB K.  
APPLICANT: SHANDILYA, HARINI  
TITLE OF INVENTION: Cloning and Expression of Rat Liver and

TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.  
 STREET: 1100 NEW YORK AVE., N.W., SUITE 600  
 CITY: WASHINGTON  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20005-3934  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/795,395  
 FILING DATE: 04-FEB-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/024,057  
 FILING DATE: 16-AUG-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: ESMOND, ROBERT W.  
 REGISTRATION NUMBER: 32, 893  
 REFERENCE/DOCKET NUMBER: 0942.3440002  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-371-2600  
 TELEFAX: 202-371-2540  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 456 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-795-395-2

Query Match 8.8%; Score 621; DB 2; Length 456;  
 Best Local Similarity 31.9%; Pred. No. 7.4e-52;  
 Matches 151; Conservative 86; Mismatches 210; Indels 26; Gaps 5;

655 NLDLIASSFCLOHCPYLKRIKRVYVKGIFPRDSABACPVVPLM---MRDKTLIEBQMED 680  
 2 NLDI-----HCQLSDAR-----WTELLPLQLQYEVVVRIDDCGLTEBHCKD 42

661 FCSMLGTHPHLRQDLGSSILTERAMKTLCAKLRHPTCKIOTLMFRNAQIT-PCVQHLMR 739  
 43 IGSALRANPSTLTELCLRTNELGDAGVHLVLOGLOSPTCKIOKLSLONCSLTEAGCGVLP 102

740 IVMANRNLRSINLIGGTHLKEEDVMACEALKHPKCLLESRLDCCGLTHACVLYKSQILT 799  
 103 TIRSLPTLRBELHLSNPLGADGRLRLCEGLDPOCHLEKQLEVCRLTPAASCEPLASVLR 162

800 TSPSLKSLSLAGNKVTQGVMPISDALRVSCALQKILIEDCGITATGCGSLASALVSNR 859  
 163 ATRALKELTVSNNDIGEGARVILGQGLADSAQQLTEIRLENGLTPANCKDLGIVASQA 222

860 SLTHCLSNNSLIGNEGVNLCRSWRLPHCSIORMLNQCHLDTAGCGFLALALMGNSWLT 919  
 223 SLRELDLGSNGIGDAGIELCPGLLSPASRLKTLMLMECDITPASCGRDLCRVLAQKETLK 282

920 HLSLSMNVENDGVNLTCEVWREPSCHQDLELVYCHLTAAACESLSCVYSRHLKSLD 979  
 283 ELSLGNKLGDEGALLCEBSLQPCQLESIMVKSCLTPAACQHVSLMTLTKNKLLELQ 342

980 LTDNAIGDGVNAAACEGKOKNSVLTGLKAKGLTSDCCBALSLATLSCNRHLTSLNVQ 1039  
 343 LSSNLTGDSGIGELQALSQPETTLRVLCIDGCEVTNNGSSSLAALLANSILRELDLSN 402

1040 NNFSPKGMKLCSPAACPSTSNLIQIGLMKWOYFVOIRKLLEEVQLKP-RVY 1090  
 403 NCVGDPVLTQLIGSLDEPGCALBQLVLVDTYVTEVEEDRLQALBESGKGLRVL 455

RESULT 10  
 US-08-910-731-8  
 Sequence 8, Application US/08910731  
 Patent No. 5932440  
 GENERAL INFORMATION:  
 APPLICANT: CHATTERJEE, DEB K.  
 APPLICANT: SHANDILYA, HARINI  
 TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof  
 NUMBER OF SEQUENCES: 16  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
 STREET: 1100 NEW YORK AVE., N.W., SUITE 600  
 CITY: WASHINGTON  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20005-3934  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/910,731  
 FILING DATE: (Herewith)  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/795,395  
 FILING DATE: 04-FEB-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/794,546  
 FILING DATE: 03-FEB-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/024,057  
 FILING DATE: 16-AUG-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: ESMOND, ROBERT W.  
 REGISTRATION NUMBER: 32, 893  
 REFERENCE/DOCKET NUMBER: 0942.3440003  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-371-2600  
 TELEFAX: 202-371-2540  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 456 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: both  
 TOPOLOGY: both  
 MOLECULE TYPE: peptide  
 US-08-910-731-8

Query Match 8.7%; Score 612; DB 2; Length 456;  
 Best Local Similarity 32.1%; Pred. No. 5.7e-51;  
 Matches 152; Conservative 81; Mismatches 214; Indels 26; Gaps 5;

655 NLDLIASSFCLOHCPYLKRIKRVYVKGIFPRDSABACPVVPLM---MRDKTLIEBQMED 680  
 2 NLDI-----HCQLSDAR-----WTELLPLQLQYEVVVRIDDCGLTEBHCKD 42

661 FCSMLGTHPHLRQDLGSSILTERAMKTLCAKLRHPTCKIOTLMFRNAQIT-PCVQHLMR 739  
 43 IGSALRANPSTLTELCLRTNELGDAGVHLVLOGLOSPTCKIOKLSLONCSLTEAGCGVLP 102

740 IVMANRNLRSINLIGGTHLKEEDVMACEALKHPKCLLESRLDCCGLTHACVLYKSQILT 799  
 103 TIRSLPTLRBELHLSNPLGADGRLRLCEGLDPOCHLEKQLEVCRLTPAASCEPLASVLR 162

800 TSPSLKSLSLAGNKVTQGVMPISDALRVSCALQKILIEDCGITATGCGSLASALVSNR 859  
 163 ATRALKELTVSNNDIGEGARVILGQGLADSAQQLTEIRLENGLTPANCKDLGIVASQA 222

860 SLTHCLSNNSLIGNEGVNLCRSWRLPHCSIORMLNQCHLDTAGCGFLALALMGNSWLT 919  
 223 SLRELDLGSNGIGDAGIELCPGLLSPASRLKTLMLMECDITPASCGRDLCRVLAQKETLK 282





Db 834 EALK-DHPSLTTLTSLAFNGISPEGGKSLAOLKONTTLTVIWLTKNELINDESAECFAEML 892  
Qy 884 RLPHCSLORLMLNCHDLPATAGGFLATLALMGNSMLTHLSMNPVE-----DNGVKLL 936  
Db 893 RV-NQTLRHLMWLIQNRITAKGTALQALRALQKNTAITIICLNGMLIKPEAKVFEKERIT 951  
Qy 937 C 937  
Db 952 C 952

RESULT 14  
US-09-340-620A-43  
; Sequence 43, Application US/09340620A  
; Patent No. 6482933  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF  
; FILE REFERENCE: 07334-124001  
; CURRENT APPLICATION NUMBER: US/09/340,620A  
; CURRENT FILING DATE: 1999-06-28  
; PRIOR APPLICATION NUMBER: US 09/245,281  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: US 09/207,359  
; PRIOR FILING DATE: 1998-12-08  
; PRIOR APPLICATION NUMBER: US 09/099,041  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: US 09/019,942  
; NUMBER OF SEQ ID NOS: 71  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 953  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-340-620A-43

Query Match 5.3%; Score 374.5; DB 4; Length 953;  
Best Local Similarity 22.7%; Pred. No. 6,2e-27;  
Matches 232; Conservative 150; Mismatches 398; Indels 241; Gaps 44;

Qy 54 NVECLA--LLHHEYGASLAWATSIISIFENNNLRTLSKARDMKISQAMEOGATPAE 111  
Db 36 NTQCLVDNLLENQVFS-----EDAEIVACPTKDKVKRILDLVOSKGEVSE 84  
Qy 112 ---TBOGISA-----MBOGATAE-----TEOGH--GGDTWDY 143  
Db 85 FFLVYLQQLLEDAVYDLRLMLSEIGFSPQLIRTKITVNTDPVSRYTQQLRHQIGRD--- 140  
Qy 144 KSHWTKFAEEDVRSFENTAADPEMQTLA--AFPSDRMGF----- 185  
Db 141 -SKRMCLYAKQEDL--LLEET-----YDITLMGLVGFNNMLSGISGLDCLLDISTGYLN 192  
Qy 186 -RPRTVLHGKSGIGKSLARRIVCWAQGGIYOGMFSYFELPVREMRKES---SVT 241  
Db 193 EHGFEVFGDAGVKGKSWLQRLQSLMASGRL--TSTAKFFHPCRMSPCKESDMLSLQ 251  
Qy 242 EFISREW-----PDGQAPTEIMSRPERLLFIIDPFDDGS--VIANDTKLCKMAEKQRP 295  
Db 252 DLRFKHFQYRQDDEEVFSFLIRPHTALFTFDGIDELHSDPDISRVSDSCPM--BPANP 310  
Qy 296 FTLLRSLLRKVLIPESFLIVTVR--DVGTEKLKSEVVSPPYLLVYRGISGEORLHL--- 348  
Db 311 LVLLANLISGRLLKAGKLTARIGVEVPROLRKVV-----LIRGSPS--HLRAYA 361  
Qy 349 ---LLERIGIGHOKTOGLRAIMNNRELLDOCVAVGSLICVALQ--LDVVGESVAPENQ 404  
Db 362 RRMPEPRAQEHLL---LQGDANPNLCGLGVPLFCMIIFRCQHPTVEGSSSQLPD 417  
Qy 405 TLTGTHAAFAV-----FQGLTRGVYRRCNLBERVV-----LKRFRMAVVEGWNRKS 452  
Db 418 CAVTLTDLVFLVTEVHLNRQPOSSSIVORNTSPAEITLPAQWRTLHALGEVAHRTDLSLF 477

Qy 453 VFDDDDLMVQGLGSESEIRALFHNMTLLPDSHCE--EYTFPHLSLODFCALYYVLEGL 510  
Db 478 VFGGEVQAQKLDQEGDLQFLR--ALPDVGPEGGQYEFPHLLIQAFPAFLVADNDKV 535  
Qy 511 IEPALCPLYEKTGRSMELKQAGPHI-----HSLMKRFLFG 547  
Db 536 STRELLRPFREWTSFG--EATSSCHSFFSPQCLGGRSLGPPDFRNKDHFOFTNLFCVG 594  
Qy 548 LVSEDEVRRPLEVLLGCPVPLGV---KOKLHWV-----SLGQOPNATTPG----- 590  
Db 595 LIAKARQKILROL---VPKAIIRRRKAL--WHLFASLSYLSKSLPRVSGGFNOVHAM 649  
Qy 591 -DTLDAPHCLEFQDQKEFVALNPSQEWLPI---NONLDIASSFCLOHCPYLKIR 645  
Db 650 PTFLLMLRCIYEOTQSGVRLAARGISADYTLKAFACGADCSALSFVLHH--FHRQLA 707  
Qy 646 VDVGKIFPRBSAECVAVLMMRDKTLIEQWEDFCSMLGTHPHLRQLDSSILTERA 705  
Db 708 LIDLNNNINDYGVQ--ELQPCFSR-----LTVIRLSVQITDTG 744  
Qy 706 MKTLCAKLHPPTCKIOTLM--FRNAQITP--GVQHLMRIVMANRLIRSLNGTILKEEDVR 763  
Db 745 KVLCEEL--TKKITVFLGYNQITDIGARYVAQILDECRLKHLKLRKARITSSGK 802  
Qy 764 MACBALKHPKCLLESRLDCCGLTHACYLKISQILTTPSPSLKSLAGKRVTDQGVPLS 823  
Db 803 CVALAVERN-----STGIVDVGMMGNQIGEGAKAFA 833  
Qy 824 DALRVSCALQKLLIEDCGITATGCGSLASALVSNLSLTHLCISNNSLGEVNLICRSM 883  
Db 834 EALK-DHPSLTTLTSLAFNGISPEGGKSLAOLKONTTLTVIWLTKNELINDESAECFAEML 892  
Qy 884 RLPHCSLORLMLNCHDLPATAGGFLATLALMGNSMLTHLSMNPVE-----DNGVKLL 936  
Db 893 RV-NQTLRHLMWLIQNRITAKGTALQALRALQKNTAITIICLNGMLIKPEAKVFEKERIT 951  
Qy 937 C 937  
Db 952 C 952

RESULT 15  
US-09-865-364-43  
; Sequence 43, Application US/09865364  
; Patent No. 6613521  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF  
; FILE REFERENCE: 07334-112001  
; CURRENT APPLICATION NUMBER: US/09/865,364  
; CURRENT FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: US 09/207,359  
; PRIOR FILING DATE: 1998-12-08  
; PRIOR APPLICATION NUMBER: US 09/099,041  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: US 09/019,942  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 953  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-865-364-43

Query Match 5.3%; Score 374.5; DB 4; Length 953;  
Best Local Similarity 22.7%; Pred. No. 6,2e-27;  
Matches 232; Conservative 150; Mismatches 398; Indels 241; Gaps 44;

Qy 54 NVECLA--LLHHEYGASLAWATSIISIFENNNLRTLSKARDMKISQAMEOGATPAE 111  
Db 36 NTQCLVDNLLENQVFS-----EDAEIVACPTKDKVKRILDLVOSKGEVSE 84

QY 112 ---TEORISQA-----MEOGATAAE-----TEBOGH---GGDTWDX 143  
DB 85 PFLYVLOQLEDAVDRLMLSEIGSPSOLIRTKTIIVNTDPVSRYTQOLRHQGRD--- 140  
QY 144 KSHVMTKFAEBEDVRSFENTAADPEMOTLAG--AFDSRWGF----- 185  
DB 141 -SKFMLCYAQKEDL--LEET-----YMDTLMGLVGFNNENIGSLGDLCDLHSTGVLAN 192  
QY 186 -RPRIVLHGKSGIGKSALARRIVLCWAQGLYQGMFSTVFLPVREMOCKES--SVT 241  
DB 193 BHGETVVFVGDAVGKSMILQRLQSLMASGRL-TSTAKFFHFRCRMFSCFKESDMLSLQ 251  
QY 242 BRISREW----PDSQAPTEIMSRPERBLFIIDGFDLGS--VLNNDTLCCKMAEKQRP 295  
DB 252 DULFPHFCYPEODPEEVSFLLRFPHTALFTPDGLDELSDPDLRVPDSCCPW-EPAP 310  
QY 296 FTLIRSLRKVLLPESFLIVTR---DVTEKLKSEVVSPLYLVRSIGSEQRILH--- 348  
DB 311 LVLLANLISGRLLKGAGKLLTARTGVEVPROLLRKV-----LARGSPS---HLRAYA 361  
QY 349 ---LIERGIGEHQKTQGLRAIMNREBLDQCQVPAGSLICVALQ-LQDVGESVAPFNQ 404  
DB 362 RMPFERTAQEHL---LQQLDANPILCSLCGVLFCWIIIFRCFHQFQVFEQSSQLPD 417  
QY 405 TLTLGHAFFV-----FHQUTPRGVVRCLNBERVV-----LKRFCMAVEGYMNRKS 452  
DB 418 CAVTLTDVFLVTEVHLNRPPSSLVQRTNRSPAETLRAGMRTLHAGEVNAHGTDKSLF 477  
QY 453 VFDGDDLMVOGLGESELRALFMMNILLPDSHC--EYTFPHLSLODFCAALYYVLEGL 510  
DB 478 VFGQEVQASKLOEGDLQGLFR--ALPVGEGOGSYEFFHLLQAFPTAFVADKV 535  
QY 511 IEPALCPLYVEKTXSMELKQAGFI-----HSLMKRFLFG 547  
DB 536 STRELLRFREWTSPG-EATSSCHSPSPFCGLGSRSLGPDPRKNKHQFTNLFPVG 594  
QY 548 LVSEVVRBPBVLGCPPLGV--KQKLLHV-----SLGQOPNATTPG----- 590  
DB 595 LLAKAROKLLRQL---VPKALLRRRKAL-WAHLFASLRSTYKSLPRVOSGGFNQVHAM 649  
QY 591 -DTLDAFCLFETODKEFVRLALNSFOEYWLPI---NONLDLIASSFLOHCPYLKIR 645  
DB 650 PTFMLMLKCIYETQOKVGRLLAARGISADYLLAFCMACSDCSALFVLHH--FHRQLA 707  
QY 646 VDVKGI FPRDESAAECPPVPLMMRDKTLIEQWEDFCSMLGTHPHLRQLDLGSILTERA 705  
DB 708 LDLNINNINDYGVQ--ELQPCFSR-----LTVIRLSVNQITDTG 744  
QY 706 MKTILCAKLRHPTCKIQTLM-FRMAQITP-GVQHLMRYMANRLASLNLGTHLKEEDVR 763  
DB 745 VKVLCEEL--TKYKIVTEFLGLYNNQITDIGARYAQIIDECEGLKHLKIGKRLITSEGGK 802  
QY 764 MACBALKBPKCLLESBLRDCGSLTHACYLKISQIILTSPSLKSLAGNKVTDGVMPLS 823  
DB 803 CVALLAVKN-----STSIYDVGMWGNQIGDEGAKAPA 833  
QY 824 DALRVSOALQKLLIEDCGITATGCSIASALVSNRSLTHCLSNNSLGNBGNLLCRSM 883  
DB 834 BALK-DHPSLITLSLAFNGISPEGKSLAQLAKONTTLIVILTKNELNDBSABCFAMTL 892  
QY 884 RLPHCSLQRLMNOCHLDTAGCGFLALALMGNMWLTHLSLNMVPE-----DNGVKLL 936  
DB 893 RV-NQTLRHLWLQIRITAKGTAOLARALQKNTAITBEICLNGNLIKPEBAKVFEENEKRII 951  
QY 937 C 937  
DB 952 C 952

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: July 19, 2005, 11:19:57 ; Search time 475 Seconds

(without alignments)  
1098.368 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 7074  
Sequence: 1 MEGDKSLTFSSYGLQWCLYVE.....DDHSGVMSLGAAGLEGVLS 1344

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1736639 seqs, 388188149 residues

Total number of hits satisfying chosen parameters: 1736639

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubppaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubppaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubppaa/US09C\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubppaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubppaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubppaa/US10F\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubppaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7074	100.0	1344	US-10-124-498-6	Sequence 6, Appl1
2	7074	100.0	1344	US-10-066-521-6	Sequence 6, Appl1
3	5756	81.4	1162	US-10-216-645-2	Sequence 2, Appl1
4	5756	81.4	1162	US-10-416-642-1	Sequence 1, Appl1
5	5741.5	81.2	1200	US-10-399-443-24	Sequence 24, Appl1
6	5741.5	81.2	1200	US-10-677-943-24	Sequence 24, Appl1
7	5741.5	81.2	1200	US-10-860-761-4	Sequence 24, Appl1
8	5683.5	72.3	1143	US-10-216-645-4	Sequence 4, Appl1
9	5113.5	46.9	1033	US-10-092-900A-348	Sequence 348, Appl1
10	3318.5	38.0	1111	US-10-407-866-96	Sequence 96, Appl1
11	2688	38.0	1111	US-10-216-645-5	Sequence 5, Appl1

	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	
	2577	2577	2577	36.4	1111	15	US-10-399-443-6	Sequence 6, Appl1																											
	13	2577	36.4	1111	15	US-10-677-943-6	Sequence 6, Appl1																												
	14	1940	27.4	385	15	US-10-399-443-2	Sequence 2, Appl1																												
	15	1940	27.4	385	15	US-10-677-943-2	Sequence 2, Appl1																												
	16	1885.5	26.7	1093	16	US-10-794-342-13	Sequence 13, Appl1																												
	17	1434	20.3	994	14	US-09-965-621-16	Sequence 16, Appl1																												
	18	1434	20.3	994	14	US-10-124-498-24	Sequence 24, Appl1																												
	19	1434	20.3	994	14	US-10-066-521-24	Sequence 24, Appl1																												
	20	1434	20.3	994	15	US-10-407-866-16	Sequence 16, Appl1																												
	21	1434	20.3	994	16	US-10-781-294-16	Sequence 16, Appl1																												
	22	1402	19.8	924	16	US-10-357-820-52	Sequence 52, Appl1																												
	23	1332.5	18.8	1035	16	US-10-794-342-14	Sequence 14, Appl1																												
	24	1330.5	18.7	1035	15	US-09-965-621-24	Sequence 24, Appl1																												
	25	1330.5	18.7	1035	15	US-10-407-866-24	Sequence 24, Appl1																												
	26	1320.5	18.7	1035	16	US-10-781-294-24	Sequence 24, Appl1																												
	27	1315.5	18.6	919	16	US-10-094-749-2718	Sequence 2718, Appl1																												
	28	1314.5	18.6	919	16	US-10-161-493-4	Sequence 4, Appl1																												
	29	1314.5	18.6	919	16	US-10-398-037-7	Sequence 7, Appl1																												
	30	1313.5	18.6	1036	16	US-10-794-342-18	Sequence 18, Appl1																												
	31	1303.5	18.4	1061	14	US-10-124-498-18	Sequence 18, Appl1																												
	32	1303.5	18.4	1061	14	US-10-066-521-18	Sequence 18, Appl1																												
	33	1301.5	18.4	1027	15	US-10-407-866-68	Sequence 68, Appl1																												
	34	1289.5	18.2	858	9	US-09-848-035-8	Sequence 8, Appl1																												
	35	1289.5	18.2	858	9	US-09-986-224-8	Sequence 8, Appl1																												
	36	1283	18.1	258	15	US-10-399-443-4	Sequence 4, Appl1																												
	37	1283	18.1	258	15	US-10-677-943-4	Sequence 4, Appl1																												
	38	1255	17.7	1032	16	US-10-882-761-33	Sequence 33, Appl1																												
	39	1255	17.7	1034	13	US-10-127-516-5	Sequence 5, Appl1																												
	40	1255	17.7	1034	13	US-10-027-629-5	Sequence 5, Appl1																												
	41	1255	17.7	1034	14	US-10-028-374-18	Sequence 18, Appl1																												
	42	1255	17.7	1034	14	US-10-132-967-5	Sequence 5, Appl1																												
	43	1255	17.7	1034	14	US-10-183-770-18	Sequence 18, Appl1																												
	44	1255	17.7	1034	15	US-10-264-9588-22	Sequence 22, Appl1																												
	45	1248.5	17.6	1062	14	US-10-239-663-43	Sequence 43, Appl1																												

## ALIGNMENTS

RESULT 1  
US-10-124-498-6  
Sequence 6, Application US/10124498  
Publication No. US20030017983A1  
GENERAL INFORMATION:  
APPLICANT: Bertin, John  
APPLICANT: Wang, Weiye  
APPLICANT: Blatcher, Maria  
TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR  
FILE REFERENCE: 07334-367001  
CURRENT APPLICATION NUMBER: US/10/124,498  
CURRENT FILING DATE: 2002-04-17  
PRIOR APPLICATION NUMBER: 10/066,521  
PRIOR FILING DATE: 2002-01-31  
PRIOR APPLICATION NUMBER: 60/318,645  
PRIOR FILING DATE: 2001-09-10  
PRIOR APPLICATION NUMBER: 60/265,231  
PRIOR FILING DATE: 2001-01-31  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 1344  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-124-498-6  
Query Match 100.0%; Score 7074; DB 14; Length 1344;  
Beet Local Similarity 100.0%; Pred. No. 0;  
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 MEGDKSLTFSSYGLQWCLYELDKKEPOTFKELKKSSSESTTCSIPQFIENAVVECTAL 60  
Db 1 MEGDKSLTFSSYGLQWCLYELDKKEPOTFKELKKSSSESTTCSIPQFIENAVVECTAL 60

```

QY 61 LHHEYYGASLAWATSIISIFENNMNLTJSEKARDPMKKSISQAMEOGATAETEOEISQA 120
D 61 LHHEYYGASLAWATSIISIFENNMNLTJSEKARDPMKKSISQAMEOGATAETEOEISQA 120
QY 121 MEQGATAETEOEGHGDWYKSHVMTKFAEEDVRSFENTADWPEMOTLAGAFDS 180
D 121 MEQGATAETEOEGHGDWYKSHVMTKFAEEDVRSFENTADWPEMOTLAGAFDS 180
QY 181 DRMGFRPTVVLHGSGIGKSALARIVLCMAOGGLYQGMFSYVFFLPVREMOBKSSV 240
D 181 DRMGFRPTVVLHGSGIGKSALARIVLCMAOGGLYQGMFSYVFFLPVREMOBKSSV 240
QY 241 TEFTSREMPDSQAPTEIMSRPERLLFTIDGFDLGSVLNNDTKLCKDMAEKOPPTLIR 300
D 241 TEFTSREMPDSQAPTEIMSRPERLLFTIDGFDLGSVLNNDTKLCKDMAEKOPPTLIR 300
QY 301 SLRKRVLLPESFLIVTVADVTEKLSKSEVSPRYLLVNGISEQRIHLLERGI GEHOXT 360
D 301 SLRKRVLLPESFLIVTVADVTEKLSKSEVSPRYLLVNGISEQRIHLLERGI GEHOXT 360
QY 361 OGLRAIMNNRELLDQCCVPVAGSLICVALQLODVVGESVAPFNQTLTGLHAAVFOHUTP 420
D 361 OGLRAIMNNRELLDQCCVPVAGSLICVALQLODVVGESVAPFNQTLTGLHAAVFOHUTP 420
QY 421 RGVVRCCLNTEBERVYLKRCFCHMAVEGVNRRKSVFEGDGLMVOGLGESELRALFHMNILLP 480
D 421 RGVVRCCLNTEBERVYLKRCFCHMAVEGVNRRKSVFEGDGLMVOGLGESELRALFHMNILLP 480
QY 481 DSHCHEEYTFPHLSIODFCALYYVLBGLIEPALCPLYVEKTRSMELKQAGFHISLW 540
D 481 DSHCHEEYTFPHLSIODFCALYYVLBGLIEPALCPLYVEKTRSMELKQAGFHISLW 540
QY 541 MKRFLFGVSDVRRPLEVLLGCPVPLGVKOKLHMWSLLGQONATPTGTLAFLHLE 600
D 541 MKRFLFGVSDVRRPLEVLLGCPVPLGVKOKLHMWSLLGQONATPTGTLAFLHLE 600
QY 601 ETODKEFRLALNFSQEWMLPINQMLDIASSFCLOHCPYLRKIRVDYKGIFFPDESAA 660
D 601 ETODKEFRLALNFSQEWMLPINQMLDIASSFCLOHCPYLRKIRVDYKGIFFPDESAA 660
QY 661 CPVVPPLMRDXTLLIEBQWEDFCSMIGTAPHRLQDLGSSILTEBAMKTLCAKLHPCTKI 720
D 661 CPVVPPLMRDXTLLIEBQWEDFCSMIGTAPHRLQDLGSSILTEBAMKTLCAKLHPCTKI 720
QY 721 QTLMEFRNAQITPGVOHLMRIYMANRNLSLNGGTHLKEEDVRAACEALKHPKCLBSLR 780
D 721 QTLMEFRNAQITPGVOHLMRIYMANRNLSLNGGTHLKEEDVRAACEALKHPKCLBSLR 780
QY 781 LDCCGLTHACYLKTSQILITSPSLKSLSLAGNKYTDQVMPLSDALRVSOQALQKLLIED 840
D 781 LDCCGLTHACYLKTSQILITSPSLKSLSLAGNKYTDQVMPLSDALRVSOQALQKLLIED 840
QY 841 CGITATGCGSILASALVSNRSLTJLCLSNNSLGNEGVNLGRSMRLPHCSIQRLMLNQH 900
D 841 CGITATGCGSILASALVSNRSLTJLCLSNNSLGNEGVNLGRSMRLPHCSIQRLMLNQH 900
QY 901 DTACCGFLATLALMGNWSLTHLSLGMNPEVDNGVLLCEVAREPSCHODLELVCHLTA 960
D 901 DTACCGFLATLALMGNWSLTHLSLGMNPEVDNGVLLCEVAREPSCHODLELVCHLTA 960
QY 961 CCESELSCVYSRSLKSLIDLTDNALGDGVAAALCEGLKQKNSVLTBLGLKACGTLSPDCE 1020
D 961 CCESELSCVYSRSLKSLIDLTDNALGDGVAAALCEGLKQKNSVLTBLGLKACGTLSPDCE 1020
QY 1021 ALSLALSCNRHLTSLNLYVNNFSPKGMKCSAPACPTSNLQIIGLWMQYVVOIRKLE 1080
D 1021 ALSLALSCNRHLTSLNLYVNNFSPKGMKCSAPACPTSNLQIIGLWMQYVVOIRKLE 1080
QY 1081 EVOLLKRRVYVDSGWSHDFDEDRHKIGLTFRLPESRAMPCALLMGNMPEOKKRVSLAGD 1140
D 1081 EVOLLKRRVYVDSGWSHDFDEDRHKIGLTFRLPESRAMPCALLMGNMPEOKKRVSLAGD 1140
QY 1141 FKSSTREFAKSLCIATANGESQVRDNNVQSSPOPAGTEHKKDKMLSVYSGAMSETALE 1200
D 1141 FKSSTREFAKSLCIATANGESQVRDNNVQSSPOPAGTEHKKDKMLSVYSGAMSETALE 1200
QY 1201 GLGSNSADHDHGGWAMSLGBLSRGLCPVLMTTAVCPGHMERLGRGCLNSADHSG 1260
D 1201 GLGSNSADHDHGGWAMSLGBLSRGLCPVLMTTAVCPGHMERLGRGCLNSADHSG 1260
QY 1261 VWSLGAAGLEGVLVNSADHSGVAMSLGAAGLEGVLVNSADHSGVAMSLGAAGLEGVL 1320
D 1261 VWSLGAAGLEGVLVNSADHSGVAMSLGAAGLEGVLVNSADHSGVAMSLGAAGLEGVL 1320
QY 1321 SNSADHSGVAMSLGAAGLEGVLVNSADHSGVAMSLGAAGLEGVLVNSADHSGVAMSLGAAGLEGVL 1344
D 1321 SNSADHSGVAMSLGAAGLEGVLVNSADHSGVAMSLGAAGLEGVLVNSADHSGVAMSLGAAGLEGVL 1344

```

```

RESULT 2
US-10-066-521-6
; Sequence 6, Application US/1006521
; Publication No. US20030027757A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Wang, Weiye
; APPLICANT: Blatcher, Maria
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR
; FILE REFERENCE: 07334-334001
; CURRENT FILING DATE: 2002-06-25
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/318,645
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1344
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-066-521-6

Query Match 100.0%; Score 7074; DB 14; Length 1344;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGDKSLTFSSYGIQWCLYELDKKEEFTFKELLKKXSSESTTCSIPOFEIENNVETLAL 60
D 1 MEGDKSLTFSSYGIQWCLYELDKKEEFTFKELLKKXSSESTTCSIPOFEIENNVETLAL 60
QY 61 LHHEYYGASLAWATSIISIFENNMNLTJSEKARDPMKKSISQAMEOGATAETEOEISQA 120
D 61 LHHEYYGASLAWATSIISIFENNMNLTJSEKARDPMKKSISQAMEOGATAETEOEISQA 120
QY 121 MEQGATAETEOEGHGDWYKSHVMTKFAEEDVRSFENTADWPEMOTLAGAFDS 180
D 121 MEQGATAETEOEGHGDWYKSHVMTKFAEEDVRSFENTADWPEMOTLAGAFDS 180
QY 181 DRMGFRPTVVLHGSGIGKSALARIVLCMAOGGLYQGMFSYVFFLPVREMOBKSSV 240
D 181 DRMGFRPTVVLHGSGIGKSALARIVLCMAOGGLYQGMFSYVFFLPVREMOBKSSV 240
QY 241 TEFTSREMPDSQAPTEIMSRPERLLFTIDGFDLGSVLNNDTKLCKDMAEKOPPTLIR 300
D 241 TEFTSREMPDSQAPTEIMSRPERLLFTIDGFDLGSVLNNDTKLCKDMAEKOPPTLIR 300
QY 301 SLRKRVLLPESFLIVTVADVTEKLSKSEVSPRYLLVNGISEQRIHLLERGI GEHOXT 360
D 301 SLRKRVLLPESFLIVTVADVTEKLSKSEVSPRYLLVNGISEQRIHLLERGI GEHOXT 360
QY 361 OGLRAIMNNRELLDQCCVPVAGSLICVALQLODVVGESVAPFNQTLTGLHAAVFOHUTP 420
D 361 OGLRAIMNNRELLDQCCVPVAGSLICVALQLODVVGESVAPFNQTLTGLHAAVFOHUTP 420

```

```

421  RGVARCINLEERVYLKRFPCMAVEGVNMRKSVFDDDLMVQGLGESELRALFHNHLLP 480
421  RGVARCINLEERVYLKRFPCMAVEGVNMRKSVFDDDLMVQGLGESELRALFHNHLLP 480
481  DSHCEYYTFFPLSLQDFCALYYVLLEGLIEBPALCPLYVEKTKSMELKQAGFIHSLM 540
481  DSHCEYYTFFPLSLQDFCALYYVLLEGLIEBPALCPLYVEKTKSMELKQAGFIHSLM 540
481  DSHCEYYTFFPLSLQDFCALYYVLLEGLIEBPALCPLYVEKTKSMELKQAGFIHSLM 540
541  MKRFLFGLVSEEDVRRLPEVLLGCPVPLGVKQKLLHWSLLGQOPNATTPGDTLDAFHCIF 600
541  MKRFLFGLVSEEDVRRLPEVLLGCPVPLGVKQKLLHWSLLGQOPNATTPGDTLDAFHCIF 600
541  MKRFLFGLVSEEDVRRLPEVLLGCPVPLGVKQKLLHWSLLGQOPNATTPGDTLDAFHCIF 600
601  ETQDEEFRLALNSFOEWLPIINQDLIASSFCLOHCYLRKIVDVXGFPPRESABA 660
601  ETQDEEFRLALNSFOEWLPIINQDLIASSFCLOHCYLRKIVDVXGFPPRESABA 660
601  ETQDEEFRLALNSFOEWLPIINQDLIASSFCLOHCYLRKIVDVXGFPPRESABA 660
661  CPVPLMNRDKTLIEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKI 720
661  CPVPLMNRDKTLIEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKI 720
661  CPVPLMNRDKTLIEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKI 720
721  QTLMRNAQITPGVQHLWRIWMANRNLISLNGTHLKEEDVRMACEALKHPKCLLESIR 780
721  QTLMRNAQITPGVQHLWRIWMANRNLISLNGTHLKEEDVRMACEALKHPKCLLESIR 780
721  QTLMRNAQITPGVQHLWRIWMANRNLISLNGTHLKEEDVRMACEALKHPKCLLESIR 780
781  LDCCGLTHACVYKISQIILTPSPKSLSLAGNKVTDQGMPLSDLRVSOCLQKLLIED 840
781  LDCCGLTHACVYKISQIILTPSPKSLSLAGNKVTDQGMPLSDLRVSOCLQKLLIED 840
781  LDCCGLTHACVYKISQIILTPSPKSLSLAGNKVTDQGMPLSDLRVSOCLQKLLIED 840
841  CGITATGOSLASALVSNRSLTHLCLSNNSLGNENGLCRSMRLPHCSLQRLMNOCHL 900
841  CGITATGOSLASALVSNRSLTHLCLSNNSLGNENGLCRSMRLPHCSLQRLMNOCHL 900
841  CGITATGOSLASALVSNRSLTHLCLSNNSLGNENGLCRSMRLPHCSLQRLMNOCHL 900
901  DTAGCGFLALAMGNSMLTHLSLSPNPEVDNGVKLLCEVMBRPSCHLQDLBLVKCHLTPA 960
901  DTAGCGFLALAMGNSMLTHLSLSPNPEVDNGVKLLCEVMBRPSCHLQDLBLVKCHLTPA 960
901  DTAGCGFLALAMGNSMLTHLSLSPNPEVDNGVKLLCEVMBRPSCHLQDLBLVKCHLTPA 960
961  CCEBSLSCVSRSRHLSKSLDLTDNALGDGVALLCEGLKQKNSVLRGLKAGLTSDCCE 1020
961  CCEBSLSCVSRSRHLSKSLDLTDNALGDGVALLCEGLKQKNSVLRGLKAGLTSDCCE 1020
961  CCEBSLSCVSRSRHLSKSLDLTDNALGDGVALLCEGLKQKNSVLRGLKAGLTSDCCE 1020
1021  ALSLALSCNRHLTSLNLYONNFSPKGMKLCGAPACPTSNLOIIGLMMKQVVOIRKLL 1080
1021  ALSLALSCNRHLTSLNLYONNFSPKGMKLCGAPACPTSNLOIIGLMMKQVVOIRKLL 1080
1021  ALSLALSCNRHLTSLNLYONNFSPKGMKLCGAPACPTSNLOIIGLMMKQVVOIRKLL 1080
1081  EVQLLKPRVVIDGWSHFSFDEDRHKIGLTFRLPESRAMPCCALLWGMNPEQKRVSLAGD 1140
1081  EVQLLKPRVVIDGWSHFSFDEDRHKIGLTFRLPESRAMPCCALLWGMNPEQKRVSLAGD 1140
1081  EVQLLKPRVVIDGWSHFSFDEDRHKIGLTFRLPESRAMPCCALLWGMNPEQKRVSLAGD 1140
1141  FKSTRFPAKSLCLATANGESQVADVNEQSSPOPMAGTEHKODKMLSVGSGAMSETAELE 1200
1141  FKSTRFPAKSLCLATANGESQVADVNEQSSPOPMAGTEHKODKMLSVGSGAMSETAELE 1200
1141  FKSTRFPAKSLCLATANGESQVADVNEQSSPOPMAGTEHKODKMLSVGSGAMSETAELE 1200
1201  GLGNSADHDGGMAMSLGRRLSRLGCPVYMTTAVCPGHERLGSRGWCINLSDHSG 1260
1201  GLGNSADHDGGMAMSLGRRLSRLGCPVYMTTAVCPGHERLGSRGWCINLSDHSG 1260
1201  GLGNSADHDGGMAMSLGRRLSRLGCPVYMTTAVCPGHERLGSRGWCINLSDHSG 1260
1261  VWSMISGAAGLBEGLVNSADHDHSGVAMSLGAAGLBEGLVNSADHDHSGVAMSLGAAGLBEGLV 1320
1261  VWSMISGAAGLBEGLVNSADHDHSGVAMSLGAAGLBEGLVNSADHDHSGVAMSLGAAGLBEGLV 1320
1261  VWSMISGAAGLBEGLVNSADHDHSGVAMSLGAAGLBEGLVNSADHDHSGVAMSLGAAGLBEGLV 1320
1321  SNSADHDHSGVAMSLGAAGLBEGLVNSADHDHSGVAMSLGAAGLBEGLVNSADHDHSGVAMSLGAAGLBEGLV 1344
1321  SNSADHDHSGVAMSLGAAGLBEGLVNSADHDHSGVAMSLGAAGLBEGLVNSADHDHSGVAMSLGAAGLBEGLV 1344
1321  SNSADHDHSGVAMSLGAAGLBEGLVNSADHDHSGVAMSLGAAGLBEGLVNSADHDHSGVAMSLGAAGLBEGLV 1344

```

```

APPLICANT: PETERS-KOTTIG, MICHAEL
APPLICANT: BECKMANN, GEORG
TITLE OF INVENTION: HUMAN MATER PROTEINS
FILE REFERENCE: SCH-1910
CURRENT APPLICATION NUMBER: US/10/216, 645
CURRENT FILING DATE: 2003-01-21
PRIOR APPLICATION NUMBER: DE 101 39 874.3 41
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 1162
TYPE: PRT
ORGANISM: Homo sapiens
US-10-216-645-2

Query Match      81.4%; Score 5756; DB 14; Length 1162;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 1102; Conservative 2; Mismatches 0; Indels 54; Gaps 1;

QY      1  MEGDKSLTFSSYGLQWCLYEIDKSEFOTFKELKKKSESTTCSIPOFEINAVECIAL 60
DB      1  MEGDKSLTFSSYGLQWCLYEIDKSEFOTFKELKKKSESTTCSIPOFEINAVECIAL 60
QY      61  LHHEYYGASLMAWTSISIFENMNLRTLSEKARDMK----- 96
DB      61  LHHEYYGASLMAWTSISIFENMNLRTLSEKARDMKSPEDPEATMTDQPSKEKVPENK 120
QY      97  -----KISQMEQEGATAETEOEISQMEQEGA 126
DB      97  -----KISQMEQEGATAETEOEISQMEQEGA 126
QY      121  YGMTKLIGVSDISDSNNKHKYVGHSSPAISQMEQEGATAETEOEISQMEQEGA 180
DB      121  YGMTKLIGVSDISDSNNKHKYVGHSSPAISQMEQEGATAETEOEISQMEQEGA 180
QY      127  TAAETEOEGGQDWDYKSHVMTKPAEEDVRSPENTADWPENQTLAAGPDSRWGFR 186
DB      127  TAAETEOEGGQDWDYKSHVMTKPAEEDVRSPENTADWPENQTLAAGPDSRWGFR 186
QY      181  TAAETEOEGGQDWDYKSHVMTKPAEEDVRSPENTADWPENQTLAAGPDSRWGFR 240
DB      181  TAAETEOEGGQDWDYKSHVMTKPAEEDVRSPENTADWPENQTLAAGPDSRWGFR 240
QY      187  PRTVILHGKSGIGSALARRIVLCMAQGLYQGMPSYVFLPVEEMORKKSSVTEFTSR 246
DB      187  PRTVILHGKSGIGSALARRIVLCMAQGLYQGMPSYVFLPVEEMORKKSSVTEFTSR 246
QY      241  PRTVILHGKSGIGSALARRIVLCMAQGLYQGMPSYVFLPVEEMORKKSSVTEFTSR 300
DB      241  PRTVILHGKSGIGSALARRIVLCMAQGLYQGMPSYVFLPVEEMORKKSSVTEFTSR 300
QY      247  EMPDSQAPVTEIMSRPERLFIIDGPDGLSVLNDTCLCDMAEKOPFTLIRSLRKV 306
DB      247  EMPDSQAPVTEIMSRPERLFIIDGPDGLSVLNDTCLCDMAEKOPFTLIRSLRKV 306
QY      301  EMPDSQAPVTEIMSRPERLFIIDGPDGLSVLNDTCLCDMAEKOPFTLIRSLRKV 360
DB      301  EMPDSQAPVTEIMSRPERLFIIDGPDGLSVLNDTCLCDMAEKOPFTLIRSLRKV 360
QY      307  LLPSEFLIVTRVDYKSEKSVSPRYLVIRGISGEORHLLERIGEHOKTQGLRAI 366
DB      307  LLPSEFLIVTRVDYKSEKSVSPRYLVIRGISGEORHLLERIGEHOKTQGLRAI 366
QY      361  LLPSEFLIVTRVDYKSEKSVSPRYLVIRGISGEORHLLERIGEHOKTQGLRAI 420
DB      361  LLPSEFLIVTRVDYKSEKSVSPRYLVIRGISGEORHLLERIGEHOKTQGLRAI 420
QY      367  MNNRELLDQCCVPVAGSLICVALQLODVGVSVAPFNQTLGLHAAFPVHOLTGRGVRR 426
DB      367  MNNRELLDQCCVPVAGSLICVALQLODVGVSVAPFNQTLGLHAAFPVHOLTGRGVRR 426
QY      421  MNNRELLDQCCVPVAGSLICVALQLODVGVSVAPFNQTLGLHAAFPVHOLTGRGVRR 480
DB      421  MNNRELLDQCCVPVAGSLICVALQLODVGVSVAPFNQTLGLHAAFPVHOLTGRGVRR 480
QY      427  CLNLEERVYLKRFPCMAVEGVNMRKSVFDDDLMVQGLGESELRALFHNHLLPDSHCE 486
DB      427  CLNLEERVYLKRFPCMAVEGVNMRKSVFDDDLMVQGLGESELRALFHNHLLPDSHCE 486
QY      481  CLNLEERVYLKRFPCMAVEGVNMRKSVFDDDLMVQGLGESELRALFHNHLLPDSHCE 540
DB      481  CLNLEERVYLKRFPCMAVEGVNMRKSVFDDDLMVQGLGESELRALFHNHLLPDSHCE 540
QY      487  YTFEFLSLQDFCALYYVLLEGLIEBPALCPLYVEKTKSMELKQAGFIHSLMNRFL 546
DB      487  YTFEFLSLQDFCALYYVLLEGLIEBPALCPLYVEKTKSMELKQAGFIHSLMNRFL 546
QY      541  YTFEFLSLQDFCALYYVLLEGLIEBPALCPLYVEKTKSMELKQAGFIHSLMNRFL 600
DB      541  YTFEFLSLQDFCALYYVLLEGLIEBPALCPLYVEKTKSMELKQAGFIHSLMNRFL 600
QY      547  GLVSEEDVRRLPEVLLGCPVPLGVKQKLLHWSLLGQOPNATTPGDTLDAFHCIFETQDE 606
DB      547  GLVSEEDVRRLPEVLLGCPVPLGVKQKLLHWSLLGQOPNATTPGDTLDAFHCIFETQDE 606
QY      601  GLVSEEDVRRLPEVLLGCPVPLGVKQKLLHWSLLGQOPNATTPGDTLDAFHCIFETQDE 660
DB      601  GLVSEEDVRRLPEVLLGCPVPLGVKQKLLHWSLLGQOPNATTPGDTLDAFHCIFETQDE 660
QY      607  FVRLALNSFOEWLPIINQDLIASSFCLOHCYLRKIVDVXGFPPRESABAECPPVPL 666
DB      607  FVRLALNSFOEWLPIINQDLIASSFCLOHCYLRKIVDVXGFPPRESABAECPPVPL 666
QY      661  FVRLALNSFOEWLPIINQDLIASSFCLOHCYLRKIVDVXGFPPRESABAECPPVPL 720
DB      661  FVRLALNSFOEWLPIINQDLIASSFCLOHCYLRKIVDVXGFPPRESABAECPPVPL 720
QY      667  WNRDKTLIEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFR 726
DB      667  WNRDKTLIEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFR 726
QY      721  WNRDKTLIEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFR 780
DB      721  WNRDKTLIEQWEDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKIQTLMFR 780
QY      727  NAQITPGVQHLWRIWMANRNLISLNGTHLKEEDVRMACEALKHPKCLLESIRLDCCGL 786
DB      727  NAQITPGVQHLWRIWMANRNLISLNGTHLKEEDVRMACEALKHPKCLLESIRLDCCGL 786

```

```

RESULT 3
US-10-216-645-2
; Sequence 2, Application US/10216645
; Publication No. US20030125282A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; APPLICANT: LESSL, MONIKA

```

Db 781 NAQITPGVQHLMRIYMANRNLRSINLGGTHLKEEDVMAACEALKHFKCLLESRLDCCGL 840  
Qy 787 THACIKISQILITSSPSIKSLIAGNKYTDGQWMLPDLALVSCALOKLLEDCGTTAT 846  
Db 841 THACIKISQILITSSPSIKSLIAGNKYTDGQWMLPDLALVSCALOKLLEDCGTTAT 900  
Qy 847 GCOSIASALVSNRSLTHLCLSNNSLGNENYLLCRSMRLPHCSLQRLMLNCHDPTAGCG 906  
Db 901 GCOSIASALVSNRSLTHLCLSNNSLGNENYLLCRSMRLPHCSLQRLMLNCHDPTAGCG 960  
Qy 907 FLALALMGNSWLTHTLSLGNPVEDNGVYLLCEVMBEPSCHLQDLLEYKCHLTAACCELS 966  
Db 961 FLALALMGNSWLTHTLSLGNPVEDNGVYLLCEVMBEPSCHLQDLLEYKCHLTAACCELS 1020  
Qy 967 CVISRSRLKSLDITDNLADGGVYALCEGLKQKNSVLTIRGLKACGLTSDCCALSLAL 1026  
Db 1021 CVISRSRLKSLDITDNLADGGVYALCEGLKQKNSVLTIRGLKACGLTSDCCALSLAL 1080  
Qy 1027 SCNRHLTSLNLYONNFPKGMKLCSAFACPTSNLQIIGLKMQYPVQIRKLLEVOULK 1086  
Db 1081 SCNRHLTSLNLYONNFPKGMKLCSAFACPTSNLQIIGLKMQYPVQIRKLLEVOULK 1140  
Qy 1087 PRVVIDGSMHSPDEDDRH 1104  
Db 1141 PRVVIDGSMHSPDEDDRY 1158

RESULT 4  
US-10-416-642-1  
; Sequence 1, Application US/10416642  
; Publication No. US20040043452A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: RAMKUMAR, Jayalakmi  
; APPLICANT: ARVIZU, Chandra  
; TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0842 PCT  
; CURRENT APPLICATION NUMBER: US/10/416,642  
; CURRENT FILING DATE: 2003-05-13  
; PRIOR APPLICATION NUMBER: 60/249,407  
; PRIOR FILING DATE: 2000-11-15  
; NUMBER OF SEQ. ID NOS: 4  
; SOFTWARE: PERL Program  
; SEQ ID NO 1  
; LENGTH: 1162  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040043452A1 7474830C01  
US-10-416-642-1

Query Match 81.4%; Score 5756; DB 15; Length 1162;  
Best Local Similarity 95.2%; Pred. No. 0;  
Matches 1102; Conservative 2; Mismatches 0; Indels 54; Gaps 1;

Qy 1 MEGKSLTFSSYGLQWCLYELDKERFQFKELKKKSESTTCSLPOEINANVECLAL 60  
Db 1 MEGKSLTFSSYGLQWCLYELDKERFQFKELKKKSESTTCSLPOEINANVECLAL 60  
Qy 61 LHHEYGASLMAWTSISIFENNNLTLSSEKARDMK----- 96  
Db 61 LHHEYGASLMAWTSISIFENNNLTLSSEKARDMK----- 96  
Qy 97 -----KISQAMEQEGATAAETEOISQAMEQGA 126  
Db 121 YGMTKLILGVSDISDNKHKYVGIHSSGFAEISQAMEQEGATAAETEOISQAMEQGA 180  
Qy 127 TAAETEOGHGQMDYKSHMTKFAEEDVRSBPENNTAAWPMQTLAAGFSDRNGFR 186  
Db 181 TAAETEOGHGQMDYKSHMTKFAEEDVRSBPENNTAAWPMQTLAAGFSDRNGFR 240

Qy 187 PRVVLHGSGIGKSLAARRIVLQWAGGLYQGMFSPVFFLPVREMRKKESSVTEFISR 246  
Db 241 PRVVLHGSGIGKSLAARRIVLQWAGGLYQGMFSPVFFLPVREMRKKESSVTEFISR 300  
Qy 247 EMPDSQAPVTEINRPERLLPIIDGPDLDGSVLANPDKLCKDAAEKQPPFTLIRSLRKV 306  
Db 301 EMPDSQAPVTEINRPERLLPIIDGPDLDGSVLANPDKLCKDAAEKQPPFTLIRSLRKV 360  
Qy 307 LLPESFLITVRDVGTEKLSEVSPRYLLVRGISEQRTHLLERIGIHEKQTKGLRAI 366  
Db 361 LLPESFLITVRDVGTEKLSEVSPRYLLVRGISEQRTHLLERIGIHEKQTKGLRAI 420  
Qy 367 MNNRELLDQCVPAVGSILCVALQLODVESVAPFNQTLTGHAAPVFFHQLTFRGVRR 426  
Db 421 MNNRELLDQCVPAVGSILCVALQLODVESVAPFNQTLTGHAAPVFFHQLTFRGVRR 480  
Qy 427 CLNLEERVVLKRFRRMAVEGVNRKSPFDDDDLMVQGLGSEELRALPHNNILLPDSHCE 486  
Db 481 CLNLEERVVLKRFRRMAVEGVNRKSPFDDDDLMVQGLGSEELRALPHNNILLPDSHCE 540  
Qy 487 YTFPHLSLODFCALYYVLEGLIEIPALCPLYVEKTKRSMELKQAGFTHSLMKRFLF 546  
Db 541 YTFPHLSLODFCALYYVLEGLIEIPALCPLYVEKTKRSMELKQAGFTHSLMKRFLF 600  
Qy 547 GLVSEVRRPLEVLLGCPVPLGVYKQLLHWVSLGQOPNAITTPBDTLDAFHCLFETQDK 606  
Db 601 GLVSEVRRPLEVLLGCPVPLGVYKQLLHWVSLGQOPNAITTPBDTLDAFHCLFETQDK 660  
Qy 607 FVRALNSFOEWLPIQONIDLASSFCLOHCPILRKIRVDYKGIFFRDSAEKCPVPL 666  
Db 661 FVRALNSFOEWLPIQONIDLASSFCLOHCPILRKIRVDYKGIFFRDSAEKCPVPL 720  
Qy 667 WMRDKTILIEQWEDFCMLGTHPHLRQDLGSSILTERAMKTLCAKLRHTCKIQTLMFR 726  
Db 721 WMRDKTILIEQWEDFCMLGTHPHLRQDLGSSILTERAMKTLCAKLRHTCKIQTLMFR 780  
Qy 727 NAQITPGVQHLMRIYMANRNLRSINLGGTHLKEEDVMAACEALKHFKCLLESRLDCCGL 786  
Db 781 NAQITPGVQHLMRIYMANRNLRSINLGGTHLKEEDVMAACEALKHFKCLLESRLDCCGL 840  
Qy 787 THACIKISQILITSSPSIKSLIAGNKYTDGQWMLPDLALVSCALOKLLEDCGTTAT 846  
Db 841 THACIKISQILITSSPSIKSLIAGNKYTDGQWMLPDLALVSCALOKLLEDCGTTAT 900  
Qy 847 GCOSIASALVSNRSLTHLCLSNNSLGNENYLLCRSMRLPHCSLQRLMLNCHDPTAGCG 906  
Db 901 GCOSIASALVSNRSLTHLCLSNNSLGNENYLLCRSMRLPHCSLQRLMLNCHDPTAGCG 960  
Qy 907 FLALALMGNSWLTHTLSLGNPVEDNGVYLLCEVMBEPSCHLQDLLEYKCHLTAACCELS 966  
Db 961 FLALALMGNSWLTHTLSLGNPVEDNGVYLLCEVMBEPSCHLQDLLEYKCHLTAACCELS 1020  
Qy 967 CVISRSRLKSLDITDNLADGGVYALCEGLKQKNSVLTIRGLKACGLTSDCCALSLAL 1026  
Db 1021 CVISRSRLKSLDITDNLADGGVYALCEGLKQKNSVLTIRGLKACGLTSDCCALSLAL 1080  
Qy 1027 SCNRHLTSLNLYONNFPKGMKLCSAFACPTSNLQIIGLKMQYPVQIRKLLEVOULK 1086  
Db 1081 SCNRHLTSLNLYONNFPKGMKLCSAFACPTSNLQIIGLKMQYPVQIRKLLEVOULK 1140  
Qy 1087 PRVVIDGSMHSPDEDDRH 1104  
Db 1141 PRVVIDGSMHSPDEDDRY 1158

RESULT 5  
US-10-399-443-24  
; Sequence 24, Application US/10399443  
; Publication No. US20040028669A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as Represented by the  
; Secretary, Department of Health & Human Services, The National Institute c  
; APPLICANT: Health

```

; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10/399,443
; PRIOR FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 1200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-399-443-24

Query Match      81.2%; Score 5741.5; DB 15; Length 1200;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 1098; Conservative 3; Mismatches 3; Indels 41; Gaps 1;

Qy 1 MEGKSLTFSSYGLQWCLYEIDKEEFOTFKELKKKSSSTTCSIPOFEINAVECLAL 60
Db 52 MEGKSLTFSSYGLQWCLYEIDKEEFOTFKELKKKSSSTTCSIPOFEINAVECLAL 111
Qy 61 LHHEYGASLAWATSSISIFENNNLRTLSEKARDMKK----- 97
Db 112 LHHEYGASLAWATSSISIFENNNLRTLSEKARDMKKSHSPEDPEATWTDGSPSEKVPGI 171
Qy 98 -----ISQAMEBEGATAETBEQSIISQAMEBEGATAETBEQSGHGD 139
Db 172 SQAVOOSATAETKEQSIISQAMEBEGATAETBEQSIISQAMEBEGATAETBEQSGHGD 231
Qy 140 TWDYKSHMTKFAEEDVRASSENTAADMPEMOTLAGAFDSRBMGFRTVVLHGKSGIG 199
Db 232 TWDYKSHMTKFAEEDVRASSENTAADMPEMOTLAGAFDSRBMGFRTVVLHGKSGIG 291
Qy 200 KSALARIVLCMAOGLYQGMFSYVFLPVREMKKSSSTTCSIFSRMPDSQAPVEIM 259
Db 292 KSALARIVLCMAOGLYQGMFSYVFLPVREMKKSSSTTCSIFSRMPDSQAPVEIM 351
Qy 260 SRPERLFTIDGFDLGSVLNNDTYLCKDMABKOPPTLISLRKVLPPESFLIVTD 319
Db 352 SRPERLFTIDGFDLGSVLNNDTYLCKDMABKOPPTLISLRKVLPPESFLIVTD 411
Qy 320 VGTETKXSEVSPRLVIRGISGEORIHLLERGI GEHOKTQGLAIINNRELLDQCVP 379
Db 412 VGTETKXSEVSPRLVIRGISGEORIHLLERGI GEHOKTQGLAIINNRELLDQCVP 471
Qy 380 AVGSILCYALQLODVVGSVAFPNQTLTGLHAAFFHOLTGRGVARRCLNTEERVVLRKF 439
Db 472 AVGSILCYALQLODVVGSVAFPNQTLTGLHAAFFHOLTGRGVARRCLNTEERVVLRKF 531
Qy 440 CRMAVEGWNKRSYFVGDGLMWQGLGSESELRLPFMMNILLPDSHCEEYTFPHLSLQDFC 499
Db 532 CRMAVEGWNKRSYFVGDGLMWQGLGSESELRLPFMMNILLPDSHCEEYTFPHLSLQDFC 591
Qy 500 AALYVVEGLETEPLACPLYEYKTKRSMELKQAGHHSIMMKRPLRELVSEEDVRRPLEV 559
Db 592 AALYVVEGLETEPLACPLYEYKTKRSMELKQAGHHSIMMKRPLRELVSEEDVRRPLEV 651
Qy 560 LLGCEVPPLGVKQKLHWVSLGQOPNATTPGDTLDAFCLFETQDKEVRLALNSFOGVW 619
Db 652 LLGCEVPPLGVKQKLHWVSLGQOPNATTPGDTLDAFCLFETQDKEVRLALNSFOGVW 711
Qy 620 LPINONDLIASSPCLQHCPLYLRKIVDVKGI FPRDESAAECVPPVLMARDKTLIEBQWE 679
Db 712 LPINONDLIASSPCLQHCPLYLRKIVDVKGI FPRDESAAECVPPVLMARDKTLIEBQWE 771
Qy 680 DFCMLGTHPHLRQDLQSSILTERAMKTLCAKLRHPCKIQTLMFRNAQTTPGVQHLMR 739

```

```

Db 772 DFCMLGTHPHLRQDLQSSILTERAMKTLCAKLRHPCKIQTLMFRNAQTTPGVQHLMR 831
Qy 740 IWMANRILRSINLGGTHKEEDVMAACEALKHPKCLLESRLDCCGLTHACYLKISQILT 799
Db 832 IWMANRILRSINLGGTHKEEDVMAACEALKHPKCLLESRLDCCGLTHACYLKISQILT 891
Qy 800 TSPBLKSLISLAGNVTVQGVWPLSDALRVSCALQKILIEBCGITTATGCSLASALVSNR 859
Db 892 TSPBLKSLISLAGNVTVQGVWPLSDALRVSCALQKILIEBCGITTATGCSLASALVSNR 951
Qy 860 SLTHCLSNNSIAGNEGVNLCSRWRLPKHSFORLMLNQCCHDPTAGCCFLATALMGNSWLT 919
Db 952 SLTHCLSNNSIAGNEGVNLCSRWRLPKHSFORLMLNQCCHDPTAGCCFLATALMGNSWLT 1011
Qy 920 HLSLSPNPVEDNGVKLLCEVWRPEPSCHLODELVKCHLTAACCSLSCVISRSRHLKSLD 979
Db 1012 HLSLSPNPVEDNGVKLLCEVWRPEPSCHLODELVKCHLTAACCSLSCVISRSRHLKSLD 1071
Qy 980 LTDNALGDGVVAALCEGLKQNSVLTTLGLKACGLTSDCCALSLASCRHLSINLVQ 1039
Db 1072 LTDNALGDGVVAALCEGLKQNSVLTTLGLKACGLTSDCCALSLASCRHLSINLVQ 1131
Qy 1040 NNFSPKGMKLCSAFACPTSNLQIIGLKMQYPVQIRKLEEVOLNPRVVIDGSMHSFD 1099
Db 1132 NNFSPKGMKLCSAFACPTSNLQIIGLKMQYPVQIRKLEEVOLNPRVVIDGSMHSFD 1191
Qy 1100 EDDRH 1104
Db 1192 EDDRY 1196

```

## RESULT 6

US-10-677-943-24

```

; Sequence 24, Application US/10677943
; Publication No. US20040072297A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America as
; APPLICANT: represented by the Secretary of the Department of Health and
; APPLICANT: Human Services
; APPLICANT: Nelson, Lawrence
; APPLICANT: Tong, Zhi-Bin
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64780
; CURRENT APPLICATION NUMBER: US/10/677,943
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US02/09776
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 1200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-677-943-24

```

```

Query Match      81.2%; Score 5741.5; DB 15; Length 1200;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 1098; Conservative 3; Mismatches 3; Indels 41; Gaps 1;

Qy 1 MEGKSLTFSSYGLQWCLYEIDKEEFOTFKELKKKSSSTTCSIPOFEINAVECLAL 60
Db 52 MEGKSLTFSSYGLQWCLYEIDKEEFOTFKELKKKSSSTTCSIPOFEINAVECLAL 111
Qy 61 LHHEYGASLAWATSSISIFENNNLRTLSEKARDMKK----- 97
Db 112 LHHEYGASLAWATSSISIFENNNLRTLSEKARDMKKSHSPEDPEATWTDGSPSEKVPGI 171
Qy 98 -----ISQAMEBEGATAETBEQSIISQAMEBEGATAETBEQSGHGD 139

```

```
Db 172 SOAVQODSATAETKEOEISQAMEQEGATAETBEOEISQAMEQEGATAETBEQHGCD 231
Qy 140 TWDKSHVMKFAEBEEDVRSPFNTADWPEMOTLAGAFSDRNGFPRRTVVLHGKSGIG 199
Db 232 TWDKSHVMKFAEBEEDVRSPFNTADWPEMOTLAGAFSDRNGFPRRTVVLHGKSGIG 291
Qy 200 KSALARRIVLCMAOGGLYQGMFSYVFPFLPVREMOQRKKESSVTEFISREWPDSQAPVTEIM 259
Db 292 KSALARRIVLCMAOGGLYQGMFSYVFPFLPVREMOQRKKESSVTEFISREWPDSQAPVTEIM 351
Qy 260 SRPERLLFIIDGFDLGSVLNNDTKLCKDMAEKOPPTLLIRSLRKVLLPESFLIVVRD 319
Db 352 SRPERLLFIIDGFDLGSVLNNDTKLCKDMAEKOPPTLLIRSLRKVLLPESFLIVVRD 411
Qy 320 VGEKLSSEVSPRYLLVRGISEGRIHLLERIGIGHOKTOGRALIMNNRELLDOCV 379
Db 412 VGEKLSSEVSPRYLLVRGISEGRIHLLERIGIGHOKTOGRALIMNNRELLDOCV 471
Qy 380 AVGSLLICVALQLODVGESVAPFNQTLTGHAAFVPHQLTFRGVRRCLNLEERVVKRF 439
Db 472 AVGSLLICVALQLODVGESVAPFNQTLTGHAAFVPHQLTFRGVRRCLNLEERVVKRF 531
Qy 440 CRMAVEGVNKRKSVFDGDDLMVOGLGSESELRALFHMNILLPDSHCEYYTFPHLSLODFC 499
Db 532 CRMAVEGVNKRKSVFDGDDLMVOGLGSESELRALFHMNILLPDSHCEYYTFPHLSLODFC 591
Qy 500 AALYYVLEGEIEIPALCPLYVEKTKRSMELKQAGPHHSIMMKRFLGLVSEEDVRRLLEV 559
Db 592 AALYYVLEGEIEIPALCPLYVEKTKRSMELKQAGPHHSIMMKRFLGLVSEEDVRRLLEV 651
Qy 560 LAGCPVPLGVKOKLLHWSVSLGQOPNATTPGDTLDAFHCLFETODKEFEVRLALNSFOEVM 619
Db 652 LAGCPVPLGVKOKLLHWSVSLGQOPNATTPGDTLDAFHCLFETODKEFEVRLALNSFOEVM 711
Qy 620 LPINQNDLILASSFCLOHCPYLKIRIVDVKGIFPPDESABACPVPLMMDKTLIEBQWE 679
Db 712 LPINQNDLILASSFCLOHCPYLKIRIVDVKGIFPPDESABACPVPLMMDKTLIEBQWE 771
Qy 680 DFCGMLGTHPHLRQLODGLSSILTERAMKTLCAKLRHPCKIOTLMPRNAQITPGVQHLMR 739
Db 772 DFCGMLGTHPHLRQLODGLSSILTERAMKTLCAKLRHPCKIOTLMPRNAQITPGVQHLMR 831
Qy 740 IVMANRNLRSNLGTHLKEEDVMAACALGHPKCLLESRLDCCGTLTHACYLKIISQILT 799
Db 832 IVMANRNLRSNLGTHLKEEDVMAACALGHPKCLLESRLDCCGTLTHACYLKIISQILT 891
Qy 800 TSPBLKSLISLAGNRYTDOGVMPISDALRVSCALQKLLIEDCGITATGCGSLASALVSNR 859
Db 892 TSPBLKSLISLAGNRYTDOGVMPISDALRVSCALQKLLIEDCGITATGCGSLASALVSNR 951
Qy 860 SLTTHCLSNNSLGNBEGVNLGRSMRBLPHCSLQRLMLNCHLDTAGCGSLALALMGNSWLT 919
Db 952 SLTTHCLSNNSLGNBEGVNLGRSMRBLPHCSLQRLMLNCHLDTAGCGSLALALMGNSWLT 1011
Qy 920 HLSTSNMPVEDNGVYKLCVWRSEPSCHLODELYKCHLTAACSESLSCVISRSHLKSID 979
Db 1012 HLSTSNMPVEDNGVYKLCVWRSEPSCHLODELYKCHLTAACSESLSCVISRSHLKSID 1071
Qy 980 LITDNLADGVAALCEGLKQKNSVLTGLKACGLTSDCCBALSLASCNHLTSLNLVQ 1039
Db 1072 LITDNLADGVAALCEGLKQKNSVLTGLKACGLTSDCCBALSLASCNHLTSLNLVQ 1131
Qy 1040 NNFBPKGMKLCSAFACPTSNLQITIGLKMQYPVQIRKLEEVOLKPRVVVIDSSWMSFD 1099
Db 1132 NNFBPKGMKLCSAFACPTSNLQITIGLKMQYPVQIRKLEEVOLKPRVVVIDSSWMSFD 1191
Qy 1100 EDDRR 1104
Db 1192 EDDRR 1196
```

```
US-10-860-761-4
; Sequence 4, Application US/10860761
; Publication No. US20040248775A1
; GENERAL INFORMATION:
; APPLICANT: WYETH
; TITLE OF INVENTION: METHODS FOR SCREENING INHIBITORS OF APOPTOSIS
; FILE REFERENCE: AM101318
; CURRENT APPLICATION NUMBER: US/10/860,761
; CURRENT FILING DATE: 2004-06-03
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 1200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-860-761-4

Query Match 81.2%; Score 5741.5; DB 16; Length 1200;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 1098; Conservative 3; Mismatches 3; Indels 41; Gaps 1;
```

```
Qy 1 MEGDKSLTFSSYGLQWCLVELDKKEPOTFKELKKXSSSTGSIPOFEIENANVECLAL 60
Db 52 MEGDKSLTFSSYGLQWCLVELDKKEPOTFKELKKXSSSTGSIPOFEIENANVECLAL 111
Qy 61 LHEHYGASLAWATSIISIFENMNLRTSEKARDMMK----- 97
Db 112 LHEHYGASLAWATSIISIFENMNLRTSEKARDMMK----- 171
Qy 98 -----ISQAMEQEGATAETBEOEISQAMEQEGATAETBEQHGCD 139
Db 172 SOAVQODSATAETKEOEISQAMEQEGATAETBEOEISQAMEQEGATAETBEQHGCD 231
Qy 140 TWDKSHVMKFAEBEEDVRSPFNTADWPEMOTLAGAFSDRNGFPRRTVVLHGKSGIG 199
Db 232 TWDKSHVMKFAEBEEDVRSPFNTADWPEMOTLAGAFSDRNGFPRRTVVLHGKSGIG 291
Qy 200 KSALARRIVLCMAOGGLYQGMFSYVFPFLPVREMOQRKKESSVTEFISREWPDSQAPVTEIM 259
Db 292 KSALARRIVLCMAOGGLYQGMFSYVFPFLPVREMOQRKKESSVTEFISREWPDSQAPVTEIM 351
Qy 260 SRPERLLFIIDGFDLGSVLNNDTKLCKDMAEKOPPTLLIRSLRKVLLPESFLIVVRD 319
Db 352 SRPERLLFIIDGFDLGSVLNNDTKLCKDMAEKOPPTLLIRSLRKVLLPESFLIVVRD 411
Qy 320 VGEKLSSEVSPRYLLVRGISEGRIHLLERIGIGHOKTOGRALIMNNRELLDOCV 379
Db 412 VGEKLSSEVSPRYLLVRGISEGRIHLLERIGIGHOKTOGRALIMNNRELLDOCV 471
Qy 380 AVGSLLICVALQLODVGESVAPFNQTLTGHAAFVPHQLTFRGVRRCLNLEERVVKRF 439
Db 472 AVGSLLICVALQLODVGESVAPFNQTLTGHAAFVPHQLTFRGVRRCLNLEERVVKRF 531
Qy 440 CRMAVEGVNKRKSVFDGDDLMVOGLGSESELRALFHMNILLPDSHCEYYTFPHLSLODFC 499
Db 532 CRMAVEGVNKRKSVFDGDDLMVOGLGSESELRALFHMNILLPDSHCEYYTFPHLSLODFC 591
Qy 500 AALYYVLEGEIEIPALCPLYVEKTKRSMELKQAGPHHSIMMKRFLGLVSEEDVRRLLEV 559
Db 592 AALYYVLEGEIEIPALCPLYVEKTKRSMELKQAGPHHSIMMKRFLGLVSEEDVRRLLEV 651
Qy 560 LAGCPVPLGVKOKLLHWSVSLGQOPNATTPGDTLDAFHCLFETODKEFEVRLALNSFOEVM 619
Db 652 LAGCPVPLGVKOKLLHWSVSLGQOPNATTPGDTLDAFHCLFETODKEFEVRLALNSFOEVM 711
Qy 620 LPINQNDLILASSFCLOHCPYLKIRIVDVKGIFPPDESABACPVPLMMDKTLIEBQWE 679
Db 712 LPINQNDLILASSFCLOHCPYLKIRIVDVKGIFPPDESABACPVPLMMDKTLIEBQWE 771
Qy 680 DFCGMLGTHPHLRQLODGLSSILTERAMKTLCAKLRHPCKIOTLMPRNAQITPGVQHLMR 739
Db 772 DFCGMLGTHPHLRQLODGLSSILTERAMKTLCAKLRHPCKIOTLMPRNAQITPGVQHLMR 831
```

```

Oy 740 |VMA|NR|NL|RS|NL|AG|STH|KE|EDV|RM|AC|AL|KH|PK|CL|ES|LR|DC|CG|LTH|AC|Y|KI|SO|LT 799
|-----|
Db 832 |VMA|NR|NL|RS|NL|AG|STH|KE|EDV|RM|AC|AL|KH|PK|CL|ES|LR|DC|CG|LTH|AC|Y|KI|SO|LT 891
Oy 800 |TSP|SL|KS|SL|AG|NK|VT|DO|G|V|MP|LSD|AL|RV|SQ|AL|OK|L|L|ED|CG|IT|AT|GC|OS|LAS|AL|V|NR 859
|-----|
Db 892 |TSP|SL|KS|SL|AG|NK|VT|DO|G|V|MP|LSD|AL|RV|SQ|AL|OK|L|L|ED|CG|IT|AT|GC|OS|LAS|AL|V|NR 951
Oy 860 |SL|TH|CL|SN|SL|GN|GV|NL|CR|SM|L|PH|CS|LO|RL|ML|NO|CH|L|D|TA|CC|SG|L|AL|AL|MG|NS|W|LT 919
|-----|
Db 952 |SL|TH|CL|SN|SL|GN|GV|NL|CR|SM|L|PH|CS|LO|RL|ML|NO|CH|L|D|TA|CC|SG|L|AL|AL|MG|NS|W|LT 1011
Oy 920 |HLS|SM|NE|VED|NG|V|L|CE|V|NR|B|PS|CH|OD|LE|V|K|CH|L|TA|AC|CE|SL|CV|IS|RS|RL|KS|LD 979
|-----|
Db 1012 |HLS|SM|NE|VED|NG|V|L|CE|V|NR|B|PS|CH|OD|LE|V|K|CH|L|TA|AC|CE|SL|CV|IS|RS|RL|KS|LD 1071
Oy 960 |LTD|NAL|GD|G|V|A|AL|CE|G|L|K|OK|NS|V|L|TR|L|GL|KA|CG|L|TS|D|CE|AL|SL|AL|SC|NR|L|TS|NL|V|Q 1039
|-----|
Db 1072 |LTD|NAL|GD|G|V|A|AL|CE|G|L|K|OK|NS|V|L|TR|L|GL|KA|CG|L|TS|D|CE|AL|SL|AL|SC|NR|L|TS|NL|V|Q 1131
Oy 1040 |N|NF|SP|KG|MM|K|C|SA|FA|CT|PS|NL|Q|IT|GL|MK|Q|Y|P|VO|IR|KL|EB|V|OL|L|K|RV|V|ID|GS|W|HS|FD 1099
|-----|
Db 1132 |N|NF|SP|KG|MM|K|C|SA|FA|CT|PS|NL|Q|IT|GL|MK|Q|Y|P|VO|IR|KL|EB|V|OL|L|K|RV|V|ID|GS|W|HS|FD 1191
Oy 1100 |ED|DR|H 1104
|-----|
Db 1192 |ED|DR|Y 1196

```

## RESULT 8

```

US-10-216-645-4
; Sequence 4, Application US/10216645
; Publication No. US20030125282A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; APPLICANT: LESSL, MONIKA
; APPLICANT: PETERS-KOTTIG, MICHAELE
; APPLICANT: BECKMANN, GEORG
; TITLE OF INVENTION: HUMAN WATER PROTEINS
; FILE REFERENCE: SCH-1910
; CURRENT APPLICATION NUMBER: US/10/216,645
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-216-645-4

```

```

Query Match      80.3%; Score 5683.5; DB 14; Length 1143;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 1087; Conservative 5; Mismatches 12; Indels 35; Gaps 2;

```

```

Oy 1 |MEGD|SL|TF|SS|Y|G|L|Q|W|C|L|Y|E|L|D|K|E|F|Q|F|K|E|L|K|K|K|S|E|S|T|G|S|I|P|Q|E|I|E|N|A|N|V|E|C|I|A|L 60
|-----|
Db 1 |MEGD|SL|TF|SS|Y|G|L|Q|W|C|L|Y|E|L|D|K|E|F|Q|F|K|E|L|K|K|K|S|E|S|T|G|S|I|P|Q|E|I|E|N|A|N|V|E|C|I|A|L 60
Oy 61 |LH|EY|Y|G|S|L|MA|W|T|S|I|S|F|E|N|N|L|T|L|S|E|K|A|D|D|K|K|I|S|Q|-----|A|ME|B|G|A|T|A|T|E|B|Q|-----| 115
|-----|
Db 61 |LH|EY|Y|G|S|L|MA|W|T|S|I|S|F|E|N|N|L|T|L|S|E|K|A|D|D|K|K|I|S|Q|-----|A|ME|B|G|A|T|A|T|E|B|Q|-----| 115
Oy 116 |-----|S|I|S|Q|A|ME|B|G|A|T|A|T|E|B|Q|G|G|D|T|W|D|Y|K|S 145
|-----|
Db 121 |Y|G|MT|L|L|I|G|V|D|I|S|D|S|N|K|K|V|T|G|H|S|F|A|B|I|S|Q|A|ME|B|G|A|T|A|T|E|B|Q|G|G|D|T|W|D|Y|K|S 180
Oy 146 |H|V|MT|F|A|E|E|D|V|R|S|F|E|N|T|A|D|W|P|E|M|Q|L|A|G|F|D|S|D|R|W|G|F|R|P|R|V|V|L|H|K|G|I|G|S|A|L|A|R 205
|-----|
Db 181 |H|V|MT|F|A|E|E|D|V|R|S|F|E|N|T|A|D|W|P|E|M|Q|L|A|G|F|D|S|D|R|W|G|F|R|P|R|V|V|L|H|K|G|I|G|S|A|L|A|R 240
Oy 206 |R|I|V|C|MA|G|G|L|Y|Q|G|N|F|S|V|F|L|P|V|E|N|Q|R|K|K|S|S|T|E|I|S|K|E|W|P|D|S|Q|A|P|V|E|I|N|S|P|R|L 265

```

```

Db 241 |R|I|V|C|MA|G|G|L|Y|Q|G|N|F|S|V|F|L|P|V|E|N|Q|R|K|K|S|S|T|E|I|S|K|E|W|P|D|S|Q|A|P|V|E|I|N|S|P|R|L 300
Oy 266 |L|P|I|D|G|F|D|I|G|S|V|I|N|D|T|K|C|K|NA|E|K|O|P|E|T|L|I|S|L|K|V|L|B|E|S|F|I|V|T|R|V|D|G|E|K|L 325
|-----|
Db 301 |L|P|I|D|G|F|D|I|G|S|V|I|N|D|T|K|C|K|NA|E|K|O|P|E|T|L|I|S|L|K|V|L|B|E|S|F|I|V|T|R|V|D|G|E|K|L 360
Oy 326 |K|S|E|V|S|P|R|Y|L|V|R|G|I|S|G|O|R|H|I|L|E|R|G|I|G|H|O|T|O|G|I|R|A|I|N|N|R|E|L|D|O|C|O|V|A|N|G|S|I 385
|-----|
Db 361 |K|S|E|V|S|P|R|Y|L|V|R|G|I|S|G|O|R|H|I|L|E|R|G|I|G|H|O|T|O|G|I|R|A|I|N|N|R|E|L|D|O|C|O|V|A|N|G|S|I 420
Oy 386 |C|V|A|L|O|D|V|E|S|V|A|P|E|N|Q|L|T|G|L|H|A|F|V|F|Q|L|P|R|G|V|R|C|L|N|E|E|V|V|L|K|R|C|R|MA|VE 445
|-----|
Db 421 |C|V|A|L|O|D|V|E|S|V|A|P|E|N|Q|L|T|G|L|H|A|F|V|F|Q|L|P|R|G|V|R|C|L|N|E|E|V|V|L|K|R|C|R|MA|VE 480
Oy 446 |G|V|N|R|K|S|V|F|D|D|I|V|O|G|L|G|E|S|E|L|R|A|L|F|H|N|I|L|L|P|D|S|H|C|E|E|Y|T|F|H|L|S|D|P|C|A|L|Y|V 505
|-----|
Db 481 |G|V|N|R|K|S|V|F|D|D|I|V|O|G|L|G|E|S|E|L|R|A|L|F|H|N|I|L|L|P|D|S|H|C|E|E|Y|T|F|H|L|S|D|P|C|A|L|Y|V 540
Oy 506 |L|E|G|E|I|E|P|A|L|C|P|L|Y|E|K|T|K|R|S|E|L|K|O|G|F|H|I|S|L|MM|R|F|L|G|V|S|E|D|R|R|P|L|E|V|L|G|P|V 565
|-----|
Db 541 |L|E|G|E|I|E|P|A|L|C|P|L|Y|E|K|T|K|R|S|E|L|K|O|G|F|H|I|S|L|MM|R|F|L|G|V|S|E|D|R|R|P|L|E|V|L|G|P|V 600
Oy 566 |P|L|G|V|K|O|K|L|H|W|S|L|L|G|O|P|N|A|T|T|P|G|D|L|D|A|F|H|C|L|P|E|T|O|D|K|E|F|R|L|A|N|S|P|O|E|W|L|P|I|N|Q 625
|-----|
Db 601 |P|L|G|V|K|O|K|L|H|W|S|L|L|G|O|P|N|A|T|T|P|G|D|L|D|A|F|H|C|L|P|E|T|O|D|K|E|F|R|L|A|N|S|P|O|E|W|L|P|I|N|Q 660
Oy 626 |L|D|L|A|S|F|C|L|O|H|C|P|Y|R|K|I|R|V|D|V|K|I|F|P|R|D|S|A|B|A|C|P|V|P|L|MM|B|D|K|L|I|E|B|O|W|E|D|F|C|S|ML 685
|-----|
Db 661 |L|D|L|A|S|F|C|L|O|H|C|P|Y|R|K|I|R|V|D|V|K|I|F|P|R|D|S|A|B|A|C|P|V|P|L|MM|B|D|K|L|I|E|B|O|W|E|D|F|C|S|ML 720
Oy 686 |G|T|H|H|L|R|Q|L|D|G|S|S|I|T|E|R|B|A|N|K|T|I|C|A|L|R|H|T|C|I|Q|T|L|P|F|N|A|O|I|T|P|G|V|O|L|H|W|I|V|MA|NR 745
|-----|
Db 721 |G|T|H|H|L|R|Q|L|D|G|S|S|I|T|E|R|B|A|N|K|T|I|C|A|L|R|H|T|C|I|Q|T|L|P|F|N|A|O|I|T|P|G|V|O|L|H|W|I|V|MA|NR 780
Oy 746 |N|L|R|S|I|N|L|G|T|H|L|K|E|D|V|RM|A|C|E|AL|KH|PK|CL|ES|LR|DC|CG|LTH|AC|Y|KI|SO|LT|TS|PS|LX 805
|-----|
Db 781 |N|L|R|S|I|N|L|G|T|H|L|K|E|D|V|RM|A|C|E|AL|KH|PK|CL|ES|LR|DC|CG|LTH|AC|Y|KI|SO|LT|TS|PS|LX 840
Oy 806 |S|L|S|L|AG|NK|VT|DO|G|V|MP|LSD|AL|RV|SQ|AL|OK|L|L|ED|CG|IT|AT|GC|OS|LAS|AL|V|NR|SL|TH|C 865
|-----|
Db 841 |S|L|S|L|AG|NK|VT|DO|G|V|MP|LSD|AL|RV|SQ|AL|OK|L|L|ED|CG|IT|AT|GC|OS|LAS|AL|V|NR|SL|TH|C 900
Oy 866 |L|S|N|S|L|AG|N|V|NL|CR|SM|L|PH|CS|LO|RL|ML|NO|CH|L|D|TA|CC|SG|L|AL|AL|MG|NS|W|L|TH|S|L|SM 925
|-----|
Db 901 |L|S|N|S|L|AG|N|V|NL|CR|SM|L|PH|CS|LO|RL|ML|NO|CH|L|D|TA|CC|SG|L|AL|AL|MG|NS|W|L|TH|S|L|SM 960
Oy 926 |N|P|V|E|D|N|G|V|K|L|C|E|V|NR|B|PS|CH|OD|LE|V|K|CH|L|TA|AC|CE|SL|CV|IS|RS|RL|KS|LD|T|ON|AL 985
|-----|
Db 961 |N|P|V|E|D|N|G|V|K|L|C|E|V|NR|B|PS|CH|OD|LE|V|K|CH|L|TA|AC|CE|SL|CV|IS|RS|RL|KS|LD|T|ON|AL 1020
Oy 986 |G|D|G|V|A|AL|C|E|G|L|K|OK|NS|V|L|TR|L|GL|KA|CG|L|TS|D|CE|AL|SL|AL|SC|NR|L|TS|NL|V|Q|N|NF|S|P|K 1045
|-----|
Db 1021 |G|D|G|V|A|AL|C|E|G|L|K|OK|NS|V|L|TR|L|GL|KA|CG|L|TS|D|CE|AL|SL|AL|SC|NR|L|TS|NL|V|Q|N|NF|S|P|K 1080
Oy 1046 |G|MM|L|C|S|A|F|A|C|P|T|S|NL|Q|IT|GL|MK|Q|Y|P|VO|IR|KL|EB|V|OL|L|K|RV|V|ID|GS|W|HS|F|E|D|D|R|H 1104
|-----|
Db 1081 |G|MM|L|C|S|A|F|A|C|P|T|S|NL|Q|IT|GL|MK|Q|Y|P|VO|IR|KL|EB|V|OL|L|K|RV|V|ID|GS|W|HS|F|E|D|D|R|Y 1139

```

## RESULT 9

```

US-10-092-900A-348
; Sequence 348, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.

```



```
APPLICANT: Ji, Weizhen
APPLICANT: Gorman, Linda
APPLICANT: Miller, Charles E.
APPLICANT: Kekuda, Ramesh
APPLICANT: Paturajan, Meera
APPLICANT: Ganguli, Esna A.
APPLICANT: Verne, Corine A.M.
APPLICANT: Guo, Xiaojia Saaba
APPLICANT: Tchernev, Velizar T.
APPLICANT: Fernandes, Elma R.
APPLICANT: Casman, Stacie J.
APPLICANT: Malyankar, Uriel M.
APPLICANT: Gerlach, Valerie
APPLICANT: Liu, Yi
APPLICANT: Anderson, David W.
APPLICANT: Spaderma, Steven K.
APPLICANT: Catterton, Elina
APPLICANT: Leite, Mario W.
APPLICANT: Zhong, Haihong
APPLICANT: Alebrock, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Bugees, Catherine E.
FILE OF INVENTION: No. US20040043382a1e1 Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-290C
CURRENT APPLICATION NUMBER: US/10/092,900A
PRIOR FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: USSN 60/274,322
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/283,675
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: USSN 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: USSN 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/274,191
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: USSN 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: USSN 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: USSN 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: USSN 60/287,424
PRIOR FILING DATE: 2001-04-30
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 768
SEQ ID NO 348
LENGTH: 1033
TYPE: PRT
ORGANISM: Homo sapiens
US-10-092-900A-348

Query Match 72.3%; Score 5113.5; DB 15; Length 1033;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 976; Conservative 7; Mismatches 18; Indels 21; Gaps 2;
```

```
Db 188 ERLFIIDGFDLGSVLNNDTKLCKDWAEEKOPFTLIRSLRKVLPESEFLIVTRDVG 247
Qy 323 EKLSEVSPRYLLVNGISGEORIHLLBERGIEHOKTOGLRALIMNNRELLDQOVAVG 382
Db 248 EKLSSEVSPRYLLVNGISGEORIHLLBERGIEHOKTOGLRALIMNNRELLDQOVAVG 307
Qy 383 SLICVALQLODVGESVAPNPQTLTGHAAFVHFQTLPRGVARCLNLEERVVLKRCRM 442
Db 308 SLICVALQLODVGESVAPNPQTLTGHAAFVHFQTLPRGVARCLNLEERVVLKRCRM 367
Qy 443 AVBGWNRKSVFPGDDLMVQGLSEBELRALFHNHILLPSHCERYTFHSLDQFCAL 502
Db 368 AVBGWNRKSVFPGDDLMVQGLSEBELRALFHNHILLPSHCERYTFHSLDQFCAL 427
Qy 503 YVVLGELEIPALCPVLVEKTKRSMELKQGFHHSIMMRFLFGLVSEVRRLLEVLLG 562
Db 428 YVVLGELEIPALCPVLVEKTKRSMELKQGFHHSIMMRFLFGLVSEVRRLLEVLLG 487
Qy 563 CPVPLGVKQKLHNVSLIGQOPNATTPGDTLDAFHCLFETQDKRFVRLALNSFOEVLPI 622
Db 488 CPVPLGVKQKLHNVSLIGQOPNATTPGDTLDAFHCLFETQDKRFVRLALNSFOEVLPI 547
Qy 623 NONIDLIASSFCIQHCPYLKIRVDVGIIPRDESAECVPLMWRDKTLIEQWEDFC 682
Db 548 NONIDLIASSFCIQHCPYLKIRVDVGIIPRDESAECVPLMWRDKTLIEQWEDFC 607
Qy 683 SMLGTHPLHQLDGLSSILTERAMKTLCAKLRHPTCKIOTLMPRNAQITPGVOHLRIYM 742
Db 608 SMLGTHPLHQLDGLSSILTERAMKTLCAKLRHPTCKIOTLMPRNAQITPGVOHLRIYM 667
Qy 743 ANRLRLSLNGTHLKEEDVRMACEALKHPKCLLESRLDCCGLTHACYLKISQILTSP 802
Db 668 ANRLRLSLNGTHLKEEDVRMACEALKHPKCLLESRLDCCGLTHACYLKISQILTSP 727
Qy 803 SLKSLSLAGKVTDOGVMPISDLRVSOCLQKYLIEDCGITATGCCSLASALVSNRSLT 862
Db 728 SLKSLSLAGKVTDOGVMPISDLRVSOCLQKYLIEDCGITATGCCSLASALVSNRSLT 787
Qy 863 HLCISNNSLNEBGNVILCRSMRLPHCSLORLMLNOCHLDRAGGFLALMGNSWLTHLS 922
Db 788 HLCISNNSLNEBGNVILCRSMRLPHCSLORLMLNOCHLDRAGGFLALMGNSWLTHLS 847
Qy 923 LSMNPVEDNGVKLLCEVWRBPSCHLQDLBYKCHLTHAACCESSLSVSRSHLKSJDLTD 982
Db 848 LSMNPVEDNGVKLLCEVWRBPSCHLQDLBYKCHLTHAACCESSLSVSRSHLKSJDLTD 907
Qy 983 NALGDGVVALCEGLKOKNSVTLRLGLKACGLTSDCEALSLALSCNRHLTSLNLVONNF 1042
Db 908 NALGDGVVALCEGLKOKNSVTLRLGLKACGLTSDCEALSLALSCNRHLTSLNLVONNF 967
Qy 1043 SPKGMMLGSAFACPTSNLQIIGLMKQYVQJRKLEEVQALKPRVVDGSMHSPFEDD 1102
Db 968 SPKGMMLGSAFACPTSNLQIIGLMKQYVQJRKLEEVQALKPRVVDGSMHSPFEDD 1027
Qy 1103 RH 1104
Db 1028 RV 1029

RESULT 10
US-10-407-866-96
; Sequence 96, Application US/10407866
; Publication No. US20040002593A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; FILE REFERENCE: 66654-10(LJ 5755)
; CURRENT APPLICATION NUMBER: US/10/407,866
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 60/370,538
; PRIOR FILING DATE: 2002-04-04
```



NUMBER OF SEQ ID NOS: 129  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 96  
 LENGTH: 682  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-407-866-96

Query Match 46.9%; Score 3318.5; DB 15; Length 682;  
 Best Local Similarity 93.8%; Pred. No. 4.9e-274;  
 Matches 640; Conservative 1; Mismatches 0; Indels 41; Gaps 1;

QY 1 MEGDKSLTFSSYGYOMCLYELDKKEFQTEKELKKSSSESTCSIPQEIENAVECAL 60  
 DB 1 MEGDKSLTFSSYGYOMCLYELDKKEFQTEKELKKSSSESTCSIPQEIENAVECAL 60  
 QY 61 LHHEYYGASLAWATSIIFENNNILRTLSKARDMKR----- 97  
 DB 61 LHHEYYGASLAWATSIIFENNNILRTLSKARDMKRSPEDPEATWTDGSPSKVPGI 120  
 QY 98 -----ISQMEQEGATAETEOEISQMEQEGATAETEOEISQMEQEGATAETEOE 139  
 DB 121 SQAVQDSATPAETKEOEISQMEQEGATAETEOEISQMEQEGATAETEOE 180  
 QY 140 TWDYKSHVMTKFAEBEDVRSFENTADMPMOTLAGAFSDRNGCFRPTVYLHGKSGIG 199  
 DB 181 TWDYKSHVMTKFAEBEDVRSFENTADMPMOTLAGAFSDRNGCFRPTVYLHGKSGIG 240  
 QY 200 KSALARRIVLCMAOGGLYQGMFSYVFLPVREMOEKSSVTEFISREMPDSQAPVTEIM 259  
 DB 241 KSALARRIVLCMAOGGLYQGMFSYVFLPVREMOEKSSVTEFISREMPDSQAPVTEIM 300  
 QY 260 SRPERLLPIIDGFDLGSVNLNDYLCMDMAEKOPPFLLIRSLRKVLPPESFLIVYRD 319  
 DB 301 SRPERLLPIIDGFDLGSVNLNDYLCMDMAEKOPPFLLIRSLRKVLPPESFLIVYRD 360  
 QY 320 VGTETKLSEVSPRYLVRSIGSEORIHLLERIGSEHQTKOGLRAIMNNELLDOCCVP 379  
 DB 361 VGTETKLSEVSPRYLVRSIGSEORIHLLERIGSEHQTKOGLRAIMNNELLDOCCVP 420  
 QY 380 AVGSLICVALQLODVGEVAPFNQTLTGHAAPFVHQLTPRGVRRCLNLEERVYLKRF 439  
 DB 421 AVGSLICVALQLODVGEVAPFNQTLTGHAAPFVHQLTPRGVRRCLNLEERVYLKRF 480  
 QY 440 CRMAVEGVWNRKSYFDGDDLVAGGISESELALTHMNTLLPDSHCEYETTFPHSLDQFC 499  
 DB 481 CRMAVEGVWNRKSYFDGDDLVAGGISESELALTHMNTLLPDSHCEYETTFPHSLDQFC 540  
 QY 500 AALYYVLEGLIEIPALCPLYVEKTKRSMELKQAGFHISLMKRFELGIVSEDEVRRPLEV 559  
 DB 541 AALYYVLEGLIEIPALCPLYVEKTKRSMELKQAGFHISLMKRFELGIVSEDEVRRPLEV 600  
 QY 560 LLAGCPVPLGVKOKLHWSLLGQOPNATTPGDTLDAFCHLETQDKEVFRALINSFOEWM 619  
 DB 601 LLAGCPVPLGVKOKLHWSLLGQOPNATTPGDTLDAFCHLETQDKEVFRALINSFOEWM 660  
 QY 620 LPINQNLDIASSFCLOHCPYL 641  
 DB 661 LPINQNLDIASSFCLOHCPYL 682

RESULT 11  
 US-10-216-645-5  
 Sequence 5, Application US/10216645  
 Publication No. US20030125282A1  
 GENERAL INFORMATION:  
 APPLICANT: WEISS, BERTRAM  
 APPLICANT: LESSL, MONIKA  
 APPLICANT: PETERS-KOTTIG, MICHAEL  
 APPLICANT: BECKMANN, GEORG  
 TITLE OF INVENTION: HUMAN MATER PROTEINS  
 FILE REFERENCE: SCH-1910  
 CURRENT APPLICATION NUMBER: US/10/216.645

CURRENT FILING DATE: 2003-01-21  
 PRIOR APPLICATION NUMBER: DE 101 39 874.3 41  
 PRIOR FILING DATE: 2001-08-10  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 5  
 LENGTH: 1111  
 TYPE: PRT  
 ORGANISM: Mus sp.  
 US-10-216-645-5

Query Match 38.0%; Score 2688; DB 14; Length 1111;  
 Best Local Similarity 51.6%; Pred. No. 8.8e-220;  
 Matches 544; Conservative 183; Mismatches 281; Indels 46; Gaps 10;

QY 74 TSISIFENNNILRTLSKARDMKRISQ-----AMEQEGATAETEE 114  
 DB 77 STWSPSENVSRALIKDQSEGEVQASERKMTSPENDSKIOKQGPQOETSFTLOSKEE 136  
 QY 115 QEISQMEQEGATAETEOEGDPTWYKSHVMTKFAEBEDVRSFENTADMPMOTL 174  
 DB 137 DEVTEA-----DKDNGGDLQDYRAHVIATKFDTSVDLH-----YDSPEMKLL 177  
 QY 175 AGAFSDRMGFRPRTVYLHGKSGIGKSAARRIVLCMAOGGLYQGMFSYVFLPVREMOE 234  
 DB 178 SDAPKPYOKTQPHITILHGRPGVKSALARSIVLCMAOGGLYQGMFSYVFLPVREMOE 236  
 QY 235 KKESSVTEFISREMPDSQAPVTEIMSRPERLLPIIDGFDLGSVNLNDYLCMDMAEK 293  
 DB 237 TEKSLAQLLAKECPDSMDLVTKIMSQPERLLPIIDGFDLDMDSVTLQDDMTLSRDMDEQ 296  
 QY 294 PPFLLIRSLRKVLPPESFLIVYRDVGTETKLSEVSPRYLVRSIGSEORIHLLERIG 353  
 DB 297 PLYLIMSRLRKALPQSPFLITTRNTGLEKLSMVVSPYILVEGASASRSQVLENI 356  
 QY 354 IGBHQTKOGLRAIMNNELLDOCCVPVAGSLICVALQLODVGEVAPFNQTLTGHAAP 413  
 DB 357 SNBSDRLOVHSHLSIENHQLDQCAPSVCSLVCEALDQKLGRCCTLPQTLTGLYATL 416  
 QY 414 VFHQLTPRGVRRCLNLEERVYLKRFMAVEGVWNRKSYFDGDDLVAGGISESELALF 473  
 DB 417 VFHQLTPRGVRRCLNLEERVYLKRFMAVEGVWNRKSYFDGDDLVAGGISESELALF 476  
 QY 474 HNMNTLLPDSH-CEYETTFPHSLDQFCALITYULGEL-ELIPALCPLYVEKTKRSMELKQ 531  
 DB 477 HNMNTLLPDSH-CEYETTFPHSLDQFCALITYULGEL-ELIPALCPLYVEKTKRSMELKQ 534  
 QY 532 AGFHISLMKRFELGIVSEDEVRRPLEVILGCPVPLGVKOKLHWSLLGQOPNATTPG 591  
 DB 535 TD-DTRLGMRFLPGLMNDILKTLEVLREYVLPVTEQDQHWVSLIAQOVNISTPMD 593  
 QY 592 TLDAFHCLFETQDKEVFRALINSFOEWMPLINQNLDIASSFCLOHCPYLRIKRVYKGI 651  
 DB 594 TLDAFHCLFETQDKEVFRALINSFOEWMPLINQNLDIASSFCLOHCPYLRIKRVYKGI 653  
 QY 652 PPRDESADACPVPLM-MRDKTLIEQWEDPCSMGLTHPHRLDLCSSITLERAKTTL 710  
 DB 654 LSVYNTLLELCPVYVQETQCKPLMEWNGFCSVTLGSLIRNKEBILDADSLISQPMKILC 713  
 QY 711 AKLHPHCKIOTMFRNAQITPGVQHLMRIVMANRNRLRSNLGTHLKEEDVMAACEALK 770  
 DB 714 LELRNQSCRIQKLPKSAEAVVSGUKHLMKLPFSNQNKLKYNLKGTPKDDMKLACALAK 773  
 QY 771 HPKCLLESRLDCCGLTHACYLKISQILITPSLSKSLSLAGNKVTDGVMPLPALRVSQ 830  
 DB 774 HKSCSVETLRDSEELITIGYEMISTILLITTHKCLSLAKNRYGVASMSISLGNALSSM 833  
 QY 831 CALQKTLLEDCGTATGCGSLASALVNSRLTHCLSNNSLGNEGVNLGRSMRLPHCSL 890  
 DB 834 CLLOKTLIDNCGILPASCHLVSLFENQNLTHCLSNNSLIGTEGVQOLCOFLRNPICAL 893  
 QY 891 ORLMLNCHLDTAGCGGLALALMGNSWLTHLSLSPNVEVDNGVYLCEVWRBSCCHLDL 950

Db 894 QRLIHNCHNIVDAYGFLAMRLANNTKLTLSLTWNPVGDGAMKLLCEALKEPTCYLOEL 953  
Qy 951 ELYKCHLTAACCSLSCVIRSRHLKSLDTDNALGDGVNALCEGLKOKNSVLTRLGLK 1010  
Db 954 ELVVCQLTQNCCEBDACTITTKHLKSLDLGNALGDGVTLTCEGLKQSSSLRRJGLG 1013  
Qy 1011 ACGLTSDCCALSLATSCNRHLSLTNLVONNFSPKGMKLCAPACPTSNLQIIGLWKQ 1070  
Db 1014 ACTLTSNCCALSLATSCNPHLNSLNTLVKDNFSTSGMLKLCAPACPTSNLQIIGLWKQ 1073  
Qy 1071 YPVOIRKLEEVQLLKPRVVIDSGMHSFDEDDRH 1104  
Db 1074 YVARVRQLEBEVEFKPHVVIDGDWYASDEDDRN 1107  
RESULT 12  
US-10-399-443-6  
Sequence 6, Application US/10399443  
Publication No. US20040028669A1  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America, as Represented by the  
APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
APPLICANT: Health  
APPLICANT: Nelson, Lawrence M.  
APPLICANT: Tong, Zhi-Bin  
APPLICANT: Nelson, Lawrence  
APPLICANT: Zhi-Bin, Tong  
TITLE OF INVENTION: Human Gene Critical to Fertility  
FILE REFERENCE: 4239-64785  
CURRENT APPLICATION NUMBER: US/10/399,443  
CURRENT FILING DATE: 2003-04-16  
PRIOR APPLICATION NUMBER: 60/241,510  
PRIOR FILING DATE: 2000-10-18  
PRIOR APPLICATION NUMBER: PCT/US01/10981  
PRIOR FILING DATE: 2001-04-04  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: Patent version 3.1  
SEQ ID NO 6  
LENGTH: 1111  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-399-443-6  
Query Match 36.4%; Score 2577; DB 15; Length 1111;  
Best Local Similarity 50.8%; Pred. No. 2.7e-210;  
Matches 535; Conservative 183; Mismatches 290; Indels 46; Gaps 10;  
Qy 74 TSISIFENMILRTLSERAKDDMKISQ-----AMEOGATPAETEE 114  
Db 77 STMSPESEVSHALIKDQSEVEEQASERKMTSPENDSKSIQKQGPQEQTSFTLQSKEE 136  
Qy 115 QEISQAMEQEGATAETEEQGHGGDTWDYKSHWNTKRAEEDVRSRENTAAMPENQTL 174  
Db 137 DEVEEA-----DKONGGDLDDYKAHVIAKTDVTDLH-----YDSPEKKLL 177  
Qy 175 AGAFDSRWGFRPRTVVLHGSGIGKSLARIVLCMAQGLYQGMFSYVFLPVRMOR 234  
Db 178 SDAKPYQKTPQPTIILHGRPGVKSALASIVLGAQGLFQKM-SFVIFSVREIKT 236  
Qy 235 KKESSVTEFSREMPDSQAPVTEIMSRPERILFTIIDGFDIGSVL-NNDTYLCQDMAKQ 293  
Db 237 TEKSSLAQLTAKECPDSTDVLTAKIMSQPERILFTIIDGDMDSVLQHDMDTLSDTQDEQ 296  
Qy 294 PPFLTSLILARKVLPESFLIVTVRDVTEKLKEVUSPRVLVVRGISGBRHLLEERG 353  
Db 297 PPTLMTSLARKALLPQSFLLITTRNGLEKLKSMVVSPLTILVEGSSASRSQVLENI 356  
Qy 354 IGEHQKQGLRAIMNNEBILDQCVPAVSGSLICVALQLODVAGSVAPFNQTLTGLAAAF 413  
Db 357 SNESDRLOVHSLLENQLOFQCAPSVCSLVCAALQLOKGLGRCRTLPCQTLTGLYATL 416  
Qy 414 VFHQLTFRGVVRCNLNEERVLKRCMAVEGVWNRKSVFDDGLMVQGLGSEELRALF 473

Db 417 VFHQLTFRBPSQASLQEOQITVLGLCMAAEGVTNRSVFYDDDLKNVSLKESELIALF 476  
Qy 474 HNMILLPDSH-CERYYPFPHLSLODFPCALYYVLEGL-EYEPALCPUYVEKTRSMELQ 531  
Db 477 HNMILLVGNHSECCYFSLHSLQDFPALYYVLEGLNQHFC-FIENQHSIMVYK 534  
Qy 532 AGFHHSIAMKRPFLFGLVSDVRRPLEVLLGCPVPLGKOKLHMVSLQOQPNATTPGD 591  
Db 535 TD-DTRLNKKRFLFGLMNDILKTELEVFPEYPIVPEEQKQHTVSLAQVNGTSPMD 593  
Qy 592 TLDAFHLCTFQDEKPEFRLALNSFOEYVLPINQNLDIASSFCLQHPYLRKTRVDYKGI 651  
Db 594 TLDAFYCLFESODEEFVGGALKRFOEVTLLINQMDKVSYSYCKHQONKALRVDTRDL 653  
Qy 652 FPRDESACPVVPLM-NRDKTLIEBQWEDFCSMLGTHPHLRDLDGSSITLTERAMTLC 710  
Db 654 LSVNLTLECPVTVQETQCKPLMETTGNFCSVLSRLKELIDGDSILSQRAMKILC 713  
Qy 711 AKLRHPYCKIQTLMFRNAQITPVOHLMRTVMANRNLRSNLGTHLKEEDVMAACEALK 770  
Db 714 LEIRNGSCRIQKLTFSKAEVVSGLKHLTKLTFNSQNLKYINTLGTPMKDDMKLACALK 773  
Qy 771 HPKCLBSRLDCCGLTHACYLKSQILTTSPSLKSLSLAGNKYTDGQVNPDLARVQ 830  
Db 774 HPKCSVETLRIDSCELTIIIGYEMISTILLSTTRIKCLSLAKNRGVGWSIAGNALSSM 833  
Qy 831 CALQKLTLEQCGITATGCGSLASALVNSRLTTLCSNNSLGNQVLLCRSMRLPHCSL 890  
Db 834 CLQKLTLDNCGLTLPASCHLVLSLFSNONTLHCSNNSLGTBVGVOQLQFLNBPICAL 893  
Qy 891 QRLIHNCHNIDTAGCGFLATALMGNSWLTSLISMNPVEDNGVYLLCEVNRBPSCHLODL 950  
Db 894 QRLIHNCHNIVDAYGFLAMRLANNTKLTLSLTWNPVGDGAMKLLCEALKEPTCYLOEL 953  
Qy 951 ELYKCHLTAACCSLSCVIRSRHLKSLDTDNALGDGVNALCEGLKOKNSVLTRLGLK 1010  
Db 954 ELVVCQLTQNCCEBDACTITTKHLKSLDLGNALGDGVTLTCEGLKQSSSLRRJGLG 1013  
Qy 1011 ACGLTSDCCALSLATSCNRHLSLTNLVONNFSPKGMKLCAPACPTSNLQIIGLWKQ 1070  
Db 1014 ACTLTSNCCALSLATSCNPHLNSLNTLVKDNFSTSGMLKLCAPACPTSNLQIIGLWKQ 1073  
Qy 1071 YPVOIRKLEEVQLLKPRVVIDSGMHSFDEDDRH 1104  
Db 1074 YVARVRQLEBEVEFKPHVVIDGDWYASDEDDRN 1107  
RESULT 13  
US-10-677-943-6  
Sequence 6, Application US/10677943  
Publication No. US20040072297A1  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America as  
APPLICANT: represented by the Secretary of the Department of Health and  
APPLICANT: Human Services  
APPLICANT: Nelson, Lawrence  
APPLICANT: Tong, Zhi-Bin  
TITLE OF INVENTION: Human Gene Critical to Fertility  
FILE REFERENCE: 4239-64790  
CURRENT APPLICATION NUMBER: US/10/677,943  
CURRENT FILING DATE: 2003-10-01  
PRIOR APPLICATION NUMBER: 60/241,510  
PRIOR FILING DATE: 2000-10-18  
PRIOR APPLICATION NUMBER: PCT/US02/09776  
PRIOR FILING DATE: 2002-03-29  
PRIOR APPLICATION NUMBER: PCT/US01/10981  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: Patent version 3.2  
SEQ ID NO 6  
LENGTH: 1111  
TYPE: PRT  
ORGANISM: Mus musculus

US-10-677-943-6

Query Match 36.4%; Score 2577; DB 15; Length 1111;  
Best Local Similarity 50.8%; Pred. No. 2.7e-210;  
Matches 535; Conservative 183; Mismatches 290; Indels 46; Gaps 10;

```
QY 74 TSISIFENMNLRTLSKARDMMKISQ-----AMEGATTAETEE 114
D 77 STMBSEBVSRAILKDSGESEVEQASERKMTSPENDSKIQKQDPEBOQSETLOSEE 136
QY 115 QEISQAMEQEBQATAETEEOGHGADTWYKSHVMTKFAEBEDVRRSFENTADWPEWOTL 174
D 137 DEVEA-----DKONGGDLQDYKAHVIKFTSVLDH-----YDSEPMKLL 177
QY 175 AGARDSBKWGFRPRVTVHSGSGICKSALARIYVLCMAGSLYQGMESYVFLPREMQR 234
D 178 SDAKPYOKTQOPHTIILHGRPGVKSALARSIVLGTQGXLPQCM-SFVIFFSVREKXT 236
QY 235 KKESSVTEFISREWPDSQAPYTEIMSRPERLLFIIDGFDLGSYV-LNNDTKLCKDMAEKQ 293
D 237 TEKSLAQILAKECPDSDLVTKIMSOPERLLFVIDGDDMDSVLQHDMDTLSDPTQEQ 296
QY 294 PPTLIRSLRKVLIPESFLIVYADVGTETKSESVSPRYLLVNGISGEORIHLLERG 353
D 297 PITYMYSLLRKALLPQSFLLITTRNTGLEKKSVMVSPYLIVGSLASRSQVLENI 356
QY 354 IGEHOKTQGLAIMNRRLLDQCVPAVSGILCVALOQDVVGSVAFPNQTLTGHAAP 413
D 357 SNESDRIQVHSLLENHQLFDQCAPSVCSLVCEALQKQKGRCTLPCTLTGLYATL 416
QY 414 VFHOLTFRGVARRCLNLEBRVVLKRCFMAVEGVNRRKSVFDGDDLMVQGLGESELRALF 473
D 417 VFHOLTFRKPSQALSQEQITLVGLCMMAEGVTMTASVYDDDKYSLKESHTLALF 476
QY 474 HNNILPDSH-CBEYTFEFLSLDQFCALYYVLEGL-EIIPALCPLYVEKTKSMELKQ 531
D 477 HNNILQVGHNSQECYVFSHSLDQFALYYVLEGEETNOHC-FIENQRSIMEYKR 534
QY 532 AGFTHISLMKRFPLGVSEBVRRLPYLLGCPVPLGVKQKLLMWSLGGQPNATTPGD 591
D 535 TD-DTRLGMRKPLFGMLNKDILKLTLELFEYPVLPYTEQKLOHTVSLIAQOVNGTSPMD 593
QY 592 TLDAFHCLETQDSEFVRLAANSFOEVLPINONDLIASFCLOHCHYLKRIYDVGI 651
D 594 TLDAFHCLESQDESEFVGALKRFOEVTLLINQKMDLVSSYCLGHCONLPAIRDKDL 653
QY 652 PPRDESABCPVPLM-WRDKTLLEBQWEDFCMSLGTPHLRDLDGSSITTEBAMKTL 710
D 654 LSVNTELECPVTVVQETQCKPLMETTGNFCSVLGRNKEJDLGDSILSQRAMKILC 713
QY 711 AKLRHPCKIQTLMFRNAQITPGVOHLMRIYMANRNLISNLGTHLKEEDVRMACEALK 770
D 714 LEIRNOSCRLOKLTFSKALVAVSGKLHKLKLFNSONLKYLMNGTTPMKDDMKLACEALK 773
QY 771 HPKCLLSRLDCCGLTACVLYLSQILTSPLSKSLSLAGNXYTDCQVMLSLALRSQ 830
D 774 HPKCSVETLRDSCELTITIGYEMISTLITRKLCSLANKRGVSMISLGNALSSSM 833
QY 831 CALOKLILDECGITATGQCSLASALVSNRSITLHCLSNNSLGNQVNLICRSMRLPHCSL 890
D 834 CLQKLLINDNGLPASPCHLLVSLFNSQNTLHCLSNNSLGTBQVQGLQCFLANPBEAL 893
QY 891 QRLMLNQCGLDTPACGFLALMLGNSWLTLSLNNPYEDNGVYLCEVMREPSCHLOL 950
D 894 QRLTINHCNIVDVAYGFLANRLANNKLTHTLSLTNNPVGDMKLLCEALKEPCYLOEL 953
QY 951 ELVYCHTLAACCSLSCVISRHLKSLDLTDNMLGQGVNALLCEGLKOKXSVLTRGLK 1010
D 954 ELVYCHTLNCCEDLACMTITTKHLSLDBLGNMNLGDBGVTLTBLGSSSSLRRLGLG 1013
QY 1011 ACGTSDCCALSLATSCNRHLSLNTLVONNFSFKGMKLCSAFACPTSNLIQGLMWKQ 1070
D 1014 ACGTSDCCALSLATSCNRHLSLNTLVONNFSFKGMKLCSAFACPTSNLIQGLMWKQ 1073
```

QY 1071 YPVOIRKLEEVOLKRPVVIDSGWSHSPDEDRH 1104  
D 1074 YPARRQLESEBVFVKPHVVIDGYASDEDRN 1107

```
RESULT 14
US-10-399-443-2
; Sequence 2, Application US/10399443
; Publication No. US20040028659A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as Represented by the
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute
; APPLICANT: Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10/399,443
; PRIOR FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-399-443-2
```

Query Match 27.4%; Score 1940; DB 15; Length 385;  
Best Local Similarity 98.7%; Pred. No. 1.4e-156;  
Matches 380; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
QY 252 QAPYTEIMSRPERLLFIIDGFDLGSYVLANDTKLCCKMAEKOPPTLIRSLIRVLLPES 311
D 1 QAPYTEIMSRPERLLFIIDGFDLGSYVLANDTKLCCKMAEKOPPTLIRSLIRVLLPES 60
QY 312 FLITYADVGEKTKSESVSPRYLLVNGISGEORIHLLERGGEHOKTQGLRAIMNRE 371
D 61 FLITYADVGEKTKSESVSPRYLLVNGISGEORIHLLERGGEHOKTQGLRAIMNRE 120
QY 372 LLDQCVPAVSGILCVALOQDVVGSVAFPNQTLTGHAAPVHOLTFRGVARRCLNLE 431
D 121 LLDQCVPAVSGILCVALOQDVVGSVAFPNQTLTGHAAPVHOLTFRGVARRCLNLE 180
QY 432 ERVVLKRCFMAVGVNRRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEYTYTF 491
D 181 ERVVLKRCFMAVGVNRRKSVFDGDDLMVQGLGESELRALFHMNILLPDSHCEYTYTF 240
QY 492 HLSIQDECCALYYLLELEIEPALCPLYVEKTKRSMELKQAGFIHSLTMKRFPLGVSE 551
D 241 HLSIQDECCALYYLLELEIEPALCPLYVEKTKRSMELKQAGFIHSLTMKRFPLGVSE 300
QY 552 DVRRPLVLLGCPVPLGVKQKLLHWSLGGQPNATTPGDTLDAFHCLFETQDEFPVRLA 611
D 301 DVRRPLVLLGCPVPLGVKQKLLHWSLGGQPNATTPGDTLDAFHCLFETQDEFPVRLA 360
QY 612 LNSFOEVLPIPNQNLDIASSFCIQ 636
D 361 LNSFOEVLPIPNQNLDIASSFCIQ 385
```

RESULT 15  
US-10-677-943-2  
; Sequence 2, Application US/10677943  
; Publication No. US20040072297A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as

```
; APPLICANT: represented by the Secretary of the Department of Health and
; APPLICANT: Human Services
; APPLICANT: Nelson, Lawrence
; APPLICANT: Tong, Zhi-Bin
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64790
; CURRENT APPLICATION NUMBER: US/10/677,943
; CURRENT FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US02/09776
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: PCT/US01/10381
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 2
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-677-943-2
```

```
Query Match 27.4%; Score 1940; DB 15; Length 385;
Best Local Similarity 98.7%; Pred. No. 1.4e-156;
Matches 380; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 252 QAPTEIMSRPERLLFTIDGDDLSVINDTKLCKDMAEKQPPPTLIRSLRKVLLPES 311
Db 1 QAPTEIMSRPERLLFTIDGDDLSVINDTKLCKDMAEKQPPPTLIRSLRKVLLPES 60
QY 312 FLIYTVADVGTETKSEVSPRYLLVRGISGEORIHLLERIGEGHOKTQGLRAIMNRE 371
Db 61 FLIYTVADVGTETKSEVSPRYLLVRGISGEORIHLLERIGEGHOKTQGLRAIMNRE 120
QY 372 LLDCCVPAVGSLLCVALLQDQVVGESVAPFNQTLTGHAAFVPHQLTPRGVRCINLE 431
Db 121 LLDCCVPAVGSLLCVALLQDQVVGESVAPFNQTLTGHAAFVPHQLTPRGVRCINLE 180
QY 432 ERVVLKRCPCMAVGVNRRKSVFPGDDLMVQGLGESELRALFHMNILLPSHCEYYTF 491
Db 181 ERVVLKRCPCMAVGVNRRKSVFPGDDLMVQGLGESELRALFHMNILLPSHCEYYTF 240
QY 492 HLSIQDFCAALYYVLEGLIEIPALCPLYVEKTKRSMELKQGFHHSIMMKRFLFGVSE 551
Db 241 HLSIQDFCAALYYVLEGLIEIPALCPLYVEKTKRSMELKQGFHHSIMMKRFLFGVSE 300
QY 552 DVRRPLEVLLGCPVPLGVKQKLLHMSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLA 611
Db 301 DVRRPLEVLLGCPVPLGVKQKLLHMSLLGQQPNATTPGDTLDAFHCLFETQDKEFVRLA 360
QY 612 LNSFOEVLPIPNQNLDIASSFCIQ 636
Db 361 LNSFOEVLPIPNQNLDIASSFCIQ 385
```

```
Search completed: July 19, 2005, 12:19:11
Job time : 479 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd

OM nucleic - nucleic search, using SW model

Run on: July 18, 2005, 22:13:46 ; Search time 17004 Seconds  
(without alignments)  
11498.283 Million cell updates/sec

```

Title:      US-10-066-521-5
Perfect score: 4035
Sequence:   1 atggaagagacaatcgct.....tcgaggsgctgtgtcttcaa 4035

```

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

word size

Total number of hits satisfying chosen parameters: 9416466

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

**Post-processing: Listing first 45 summaries**

Database : GenEmbl : \*

```

1:  gb_da: *
2:  gb_htcg: *
3:  gb_in: *
4:  gb_om: *
5:  gb_ov: *
6:  gb_pat: *
7:  gb_ph: *
8:  gb_pl: *
9:  gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *

```

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2917	72.3	3489	6	AX478549	AX478549 Sequence
2	2917	72.3	3885	6	AY154460	AY154460 Homo sapi
3	2917	72.3	3926	6	AX704821	AX704821 Sequence
4	2862	70.9	3830	6	AX704823	AX704823 Sequence
5	2803	69.5	5859	6	AX459881	AX459881 Sequence
6	2803	69.5	6939	6	AX459873	AX459873 Sequence
7	2803	69.5	6939	6	AX459891	AX459891 Sequence
8	2713	67.2	3885	6	AY054986	AY054986 Homo sapi
9	2713	67.2	3900	6	AX427610	AX427610 Sequence
10	1936	48.0	2753	6	CO731113	CO731113 Sequence
11	1549	38.4	15141	9	AC011470	AC011470 Homo sapi
12	1549	38.4	19560	9	AC028480	AC028480 Homo sapi
13	1102	27.3	1157	6	AX427588	AX427588 Sequence
14	900	22.3	167509	2	AC012107	AC012107 Homo sapi
15	657	16.3	1075	6	AX427590	AX427590 Sequence
16	509	12.6	167509	2	AC012107	AC012107 Homo sapi
17	219	5.4	164824	2	AC023887	AC023887 Homo sapi
18	174	4.3	164824	2	AC023887	AC023887 Homo sapi
19	165	4.1	1705	9	BC040925	BC040925 Homo sapi

C	20	24	0.6	24	6	AX427601	Sequence
C	21	24	0.6	AX427601	4	AX427601	Sequence
C	22	23	0.6	439	4	BT00COPF	Z86019 B. taurus mr
C	23	23	0.6	3545	4	AY721594	AY721594 Bos taurus
C	24	23	0.6	159506	9	HS3418	AL021918 Human DNA
C	25	23	0.6	247877	2	AC096265	AC096265 Rattus no
C	26	23	0.6	307855	2	AC134117	AC134117 Rattus no
C	27	22	0.5	22	6	AX427594	Sequence
C	28	22	0.5	26	6	AX427596	Sequence
C	29	22	0.5	27	6	AX427604	Sequence
C	30	22	0.5	589	11	BV161832	BV161832 RPAME80
C	31	22	0.5	1576	10	F14355S07	AF143565 Mus muscu
C	32	22	0.5	3303	10	AY329488	AY329488 Mus muscu
C	33	22	0.5	3303	10	AY329491	AY329491 Mus muscu
C	34	22	0.5	3372	10	AY329485	AY329485 Mus muscu
C	35	22	0.5	3375	10	AY329489	AY329489 Mus muscu
C	36	22	0.5	3405	10	AY329487	AY329487 Mus muscu
C	37	22	0.5	3432	10	AY329484	AY329484 Mus muscu
C	38	22	0.5	3447	6	AX427592	Sequence
C	39	22	0.5	3447	10	AF074018	AF074018 Mus muscu
C	40	22	0.5	3453	10	AY329486	AY329486 Mus muscu
C	41	22	0.5	3454	10	AY329490	AY329490 Mus muscu
C	42	22	0.5	3470	10	AY196362	AY196362 Mus muscu
C	43	22	0.5	3480	10	AY196361	AY196361 Mus muscu
C	44	22	0.5	3534	10	BC053384	BC053384 Mus muscu
C	45	22	0.5	55310	2	AC102436	AC102436 Mus muscu

## ALIGNMENTS

LOCUS	AX478549	3489 bp	DNA	linear	PAT 12-AUG-2002
DEFINITION	Sequence 3 from Patent WO0246362.				
ACCESSION	AX478549				
VERSION	AX478549.1	GI:22217318			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	Ramkumar, J. and Arvizu, C.				
TITLE	Embryogenesis associated proteins				
JOURNAL	Patent: WO 0246362-A 3 20-JUN-2002; Incyte Genomics, Inc. (US)				
FEATURES	Location/Qualifiers				
source	1..3489				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
	/note="Incyte ID No: 7474830CB1"				
ORIGIN					
Query Match	72.3%	Score 2917	DB 6	Length 3489	
Best Local Similarity	99.9%	Pred. No. 0			
Matches 3017	Conservative	0	Mismatches 2	Indels 0	Gaps 0
Dy	290	AAATTTCAAGCTATGGAACAAGAGGTCCACAGCAGCAGAGACAGAAACAAGAAA	349		
Db	452	AAATTTCAAGCTATGGAACAAGAGGTCCACAGCAGCAGAGACAGAAACAAGAAA	511		
Dy	350	TTTCAACAAGCTATGGAACAAGAGGTCCACAGCAGCAGAGACAGAAACAAGACATG	409		
Db	512	TTTCAACAAGCTATGGAACAAGAGGTCCACAGCAGCAGAGACAGAAACAAGACATG	571		
Dy	410	GAGGTGACATGAGGACTACAAGAGTCAAGTGTGATGACCAATTCGTCGAGAGAGAGATG	469		
Db	572	GAGGTGACATGAGGACTACAAGAGTCAAGTGTGATGACCAATTCGTCGAGAGAGAGATG	631		
Dy	470	TACGTCGATGTTTGAACAACCTGCTGCTGACTGCGCGAATGCAAACTTGGCTGCTG	529		

## ORIGIN

	Query Match	Best Local Similarity	Matches	Conservative	Score 2917;	DB	Length 3489;	Pred. No. 0;	Mismatches	Indels	Gaps
QY	290	AAATTTACAA	GCTATGGAACAA	GGAAGGTG	CCACAGACG	CAGAGACAA	GAAACAA	GA	349		
Db	452	AAATTTACAA	GCTATGGAACAA	GGAAGGTG	CCACAGACG	CAGAGACAA	GAAACAA	GA	511		
QY	350	TTTCACAA	GCTATGGAACAA	GGAAGGTG	CCACAGACG	CAGAGACAA	GAAACAA	GA	409		
Db	512	TTTCACAA	GCTATGGAACAA	GGAAGGTG	CCACAGACG	CAGAGACAA	GAAACAA	GA	571		
QY	410	GAGGTGAC	CATATGGACTCA	AAAGTCA	CGGTATGACAA	ATTGCTGAGAGG	AGGATG		469		
Db	572	GAGGTGAC	CATATGGACTCA	AAAGTCA	CGGTATGACAA	ATTGCTGAGAGG	AGGATG		631		
QY	470	TACGTGTA	GTTTTGAAAAC	ACTGCTGCTG	ACTGCGGAAAT	GCACAACTTTG	GGCTGCTG		529		

Db 632 TACGTCGATGTTTAAAACACTGCTGTAAGTCGCGGAAATGCAAACTGTGCTG 691  
Qy 530 CTTTGTATTCAGACCGGTGGGGCTCCGGCCCTCGACGGGTGTGTGACGGAAATGAG 589  
Db 692 CTTTGTATTCAGACCGGTGGGGCTTCGGCCCTCGACGGGTGTGTGACGGAAATGAG 751  
Qy 590 GAATTTGGGAAATCGGCTCTAGCCAGAGAGATCTGTCTGTGGCGGCAAGGTGACTCT 649  
Db 752 GAATTTGGGAAATCGGCTCTAGCCAGAGAGATCTGTCTGTGGCGGCAAGGTGACTCT 811  
Qy 650 ACCAGGAAATGTTCTCTACGTCCTTCTCTCCCGTTAGAGAGATGACGGGAAAGAG 709  
Db 812 ACCAGGAAATGTTCTCTACGTCCTTCTCTCCCGTTAGAGAGATGACGGGAAAGAG 871  
Qy 710 AGAGAGGTGACAGAGTTCATCTCAGGAGGTGGCCGACTCCAGGCTCCGGTGAAG 769  
Db 872 AGAGAGGTGACAGAGTTCATCTCAGGAGGTGGCCGACTCCAGGCTCCGGTGAAG 931  
Qy 770 AGATCATGTCCCGACAGAAAGGCTGTTTTCATGATGACGGTTTCGATGACTGGGCT 829  
Db 932 AGATCATGTCCCGACAGAAAGGCTGTTTTCATGATGACGGTTTCGATGACTGGGCT 991  
Qy 830 CTGTCTCTCAACATGACAAAGGCTCTGCAAAAGACTGGGCTGAGAGAGCCTCCGTTCA 889  
Db 992 CTGTCTCTCAACATGACAAAGGCTCTGCAAAAGACTGGGCTGAGAGAGCCTCCGTTCA 1051  
Qy 890 CCTCATAGCGAGTCTGCTGAGAGAGGTCTGCTCCGAGTCTTCCGATCTGCAAG 949  
Db 1052 CCTCATAGCGAGTCTGCTGAGAGAGGTCTGCTCCGAGTCTTCCGATCTGCAAG 1111  
Qy 950 TCAGAGAGGTGGGACAGAGAGCTCAAGTCAGAGGTGCTGCTCCCGTTACCTGTTAG 1009  
Db 1112 TCAGAGAGGTGGGACAGAGAGCTCAAGTCAGAGGTGCTGCTCCCGTTACCTGTTAG 1171  
Qy 1010 TTAAGAGATCTCCGGGGAAACAAAGATCCACTTCTCTTGAAGCGGGATGTGTAGC 1069  
Db 1172 TTAAGAGATCTCCGGGGAAACAAAGATCCACTTCTCTTGAAGCGGGATGTGTAGC 1231  
Qy 1070 ATCAAGAGACAAAGGTTGCGTGCATGAAACAACGTGAGCTGCTCCGACGATGCC 1129  
Db 1232 ATCAAGAGACAAAGGTTGCGTGCATGAAACAACGTGAGCTGCTCCGACGATGCC 1291  
Qy 1130 AGGTCCCGCGGTGGGCTCTCATCTGCTGCGCTCCGAGCTGAGAGAGCTGTGGGGG 1189  
Db 1292 AGGTCCCGCGGTGGGCTCTCATCTGCTGCGCTCCGAGCTGAGAGAGCTGTGGGGG 1351  
Qy 1190 AGAGGCTGCGCCCTTCAACAAAGCTCAAGGCTTGACGCGCTTTGTGTTCAATC 1249  
Db 1352 AGAGGCTGCGCCCTTCAACAAAGCTCAAGGCTTGACGCGCTTTGTGTTCAATC 1411  
Qy 1250 AGTCACCCCTGAGGCGGTGCGGCGCTGTCTCAATCTGGAGGAAAGATGTCCTGA 1309  
Db 1412 AGTCACCCCTGAGGCGGTGCGGCGCTGTCTCAATCTGGAGGAAAGATGTCCTGA 1471  
Qy 1310 AGCGCTTCTGCGATGCTGTGAGAGAGTGTGAAATAGAAATCAAGTGTGATGATG 1369  
Db 1472 AGCGCTTCTGCGATGCTGTGAGAGAGTGTGAAATAGAAATCAAGTGTGATGATG 1531  
Qy 1370 ACAGACTCATGTTCAAGAGACTCGGGAGTCTGAGCTCCGTCTCTGTTCAATGAACA 1429  
Db 1532 ACAGACTCATGTTCAAGAGACTCGGGAGTCTGAGCTCCGTCTCTGTTCAATGAACA 1591  
Qy 1430 TCCCTTCCCGAGAGCAGCTGTGAGAGTACTACCTTCTTCAACCTCAAGTCCGAG 1489  
Db 1592 TCCCTTCCCGAGAGCAGCTGTGAGAGTACTACCTTCTTCAACCTCAAGTCCGAG 1651  
Qy 1490 ACTTCTGTGCGGCTTGTACTACGTGTTAGAGGCGCTGAAATCGAGCAGCTCTCTCC 1549  
Db 1652 ACTTCTGTGCGGCTTGTACTACGTGTTAGAGGCGCTGAAATCGAGCAGCTCTCTCC 1711  
Qy 1550 CTCTGTAGTGTGAGAGCAAAAGGCTCATGAGACTTAAACAGGACGCTTCCATATCC 1609  
Db 1712 CTCTGTAGTGTGAGAGCAAAAGGCTCATGAGACTTAAACAGGACGCTTCCATATCC 1771

Qy 1610 ACTGCTTTGGATGAAGGTTTCTTGTGCGCTCGTGAAGGAAAGATGAGAGCCAC 1669  
Db 1772 ACTGCTTTGGATGAAGGTTTCTTGTGCGCTCGTGAAGGAAAGATGAGAGCCAC 1831  
Qy 1670 TGAAGTCTCTGCGGCTGTCCGCTCCCTCGGGGGTGAAGCAGAGCTTCTGACTGG 1729  
Db 1832 TGAAGTCTCTGCGGCTGTCCGCTCCCTCGGGGGTGAAGCAGAGCTTCTGACTGG 1891  
Qy 1730 TCTCTGTTGGGTACAGGCTAATGCAACCAAGGAGAGACCTGGAGCGCTTCC 1789  
Db 1892 TCTCTGTTGGGTACAGGCTAATGCAACCAAGGAGAGACCTGGAGCGCTTCC 1951  
Qy 1790 ACTGCTTTGAGACTCAAGACAAAGATTGTTGCTTGGACTTAAACAGTTCACAG 1849  
Db 1952 ACTGCTTTGAGACTCAAGACAAAGATTGTTGCTTGGACTTAAACAGTTCACAG 2011  
Qy 1850 AAGTGTGCTTCGATTAACAGAACTGGAATTGATAGCATCTTCTGCTCCAGC 1909  
Db 2012 AAGTGTGCTTCGATTAACAGAACTGGAATTGATAGCATCTTCTGCTCCAGC 2071  
Qy 1910 ACTGTCCTATTTGCGGAAAAATGGGTGATGTCAAAGGATCTTCCAAAGATGAGT 1969  
Db 2072 ACTGTCCTATTTGCGGAAAAATGGGTGATGTCAAAGGATCTTCCAAAGATGAGT 2131  
Qy 1970 CCGCTGAGGCAATGTCTGTGGTCCCTCATGATGATGCGGATTAAGACCTCATGAGAGC 2029  
Db 2132 CCGCTGAGGCAATGTCTGTGGTCCCTCATGATGATGCGGATTAAGACCTCATGAGAGC 2191  
Qy 2030 AGTGGGAAATTTCTGCTCCATGCTTGGACCCACCCACACTGTCGGGAGCTGG 2089  
Db 2192 AGTGGGAAATTTCTGCTCCATGCTTGGACCCACCCACACTGTCGGGAGCTGG 2251  
Qy 2090 GCACAGATCTCTACAGAGCGGGCCATGAAGACCTGTGCGCAAGCTGAGGATCCA 2149  
Db 2252 GCACAGATCTCTACAGAGCGGGCCATGAAGACCTGTGCGCAAGCTGAGGATCCA 2311  
Qy 2150 CCTGCAAGATACAGACCTGATGTTTGAAGATGACAGATTACCCCTGTGTGAGACCC 2209  
Db 2312 CCTGCAAGATACAGACCTGATGTTTGAAGATGACAGATTACCCCTGTGTGAGACCC 2371  
Qy 2210 TCTGGAAGATGTCATGCGCAACCGTAACTTAAGATCTTCACTTGGAGGACCCAC 2269  
Db 2372 TCTGGAAGATGTCATGCGCAACCGTAACTTAAGATCTTCACTTGGAGGACCCAC 2431  
Qy 2270 TGAAGAAAGAGATGAAGATGCGGTGTAAGGCTTAAACACCCAAATGTTGTGG 2329  
Db 2432 TGAAGAAAGAGATGAAGATGCGGTGTAAGGCTTAAACACCCAAATGTTGTGG 2491  
Qy 2330 AGTCTTGAAGGCTGATTTGCTGTGATGACCCATGCTGTTACTGAAGATCTCCAAA 2389  
Db 2492 AGTCTTGAAGGCTGATTTGCTGTGATGACCCATGCTGTTACTGAAGATCTCCAAA 2551  
Qy 2390 TCTTTACGACTCCCGACGCTGAATCTGTAGCTTGGCAGGAAACAGGTGACAGCC 2449  
Db 2552 TCTTTACGACTCCCGACGCTGAATCTGTAGCTTGGCAGGAAACAGGTGACAGCC 2611  
Qy 2450 AGGAGTAATGCTCTAGTGAATGCTTGAAGATCTCCAGTGGCCCTGAGAAAGCTGA 2509  
Db 2612 AGGAGTAATGCTCTAGTGAATGCTTGAAGATCTCCAGTGGCCCTGAGAAAGCTGA 2671  
Qy 2510 TACTGAGAGCTGAGCATCAAGCAGGTTTCCAGAGCTTGGCTCAGCCCTCGTCA 2569  
Db 2672 TACTGAGAGCTGAGCATCAAGCAGGTTTCCAGAGCTTGGCTCAGCCCTCGTCA 2731  
Qy 2570 GCAACCGGAGCTTGAACACTGTGTCTATCAACAAAGCTTGGGAAACGAGGTGAA 2629  
Db 2732 GCAACCGGAGCTTGAACACTGTGTCTATCAACAAAGCTTGGGAAACGAGGTGAA 2791  
Qy 2630 ATCTACTGTGTGATCCATGAGGCTTCCCACTGTATCTGCAAGGCTGATGCTGAATC 2689  
Db 2792 ATCTACTGTGTGATCCATGAGGCTTCCCACTGTGTGCTGCAAGGCTGATGCTGAATC 2851

QY	2690	AGTGCACCTGGACACGGCTGGCTGTGATTTCTTGACCTGGCTTAATGGGTAACTCAT	2743
Db	2852	AGTGCACCTGGACACGGCTGGCTGTGATTTCTTGACCTGGCTTAATGGGTAACTCAT	2911
QY	2750	GGCTGACCCACCTGAGCCTTAGCATGAACCCCTGTGGAACAATATGACCTGATTCGT	2809
Db	2912	GGCTGACCCACCTGAGCCTTAGCATGAACCCCTGTGGAACAATATGACCTGATTCGT	2971
QY	2810	GGAGGTCATGAGAGAACCATCTTGTCACTTCACAGACCTGAGTTGGTAAAGTGCATC	2869
Db	2972	GCGAGGTCATGAGAGAACCATCTTGTCACTTCACAGACCTGAGTTGGTAAAGTGCATC	3031
QY	2870	TCACCCGCGCGCTGTGAGAGTCTGTCTGTGTGATCTCGAGAGACGAACACTGAAGA	2929
Db	3032	TCACCCGCGCGCTGTGAGAGTCTGTCTGTGTGATCTCGAGAGACGAACACTGAAGA	3091
QY	2930	GCTTGATCTCAACGGAACAATCCCTGGGTGACGGGTGGGTGGCTGTGCGCTGTGCGAGGAC	2989
Db	3092	GCTTGATCTCAACGGAACAATCCCTGGGTGACGGGTGGGTGGCTGTGCGCTGTGCGAGGAC	3151
QY	2990	TGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCAATGTGACTGATCTTG	3049
Db	3152	TGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCAATGTGACTGATCTTG	3211
QY	3050	ATTGTGTGAGAGCACTCTCTTGGCCCTTTCTCTGCAACCGGATCTGACCACTTAACC	3109
Db	3212	ATTGTGTGAGAGCACTCTCTTGGCCCTTTCTCTGCAACCGGATCTGACCACTTAACC	3271
QY	3110	TGTGTCAATTAATCTTCAGTCCCAAGGAATGATGAAGCTGTGTGGCCCTTGGCTGTGTC	3169
Db	3272	TGTGTCAATTAATCTTCAGTCCCAAGGAATGATGAAGCTGTGTGGCCCTTGGCTGTGTC	3331
QY	3170	CCAGCTTAATTAATGAGATTTGGGCTGTGGAATGTGCACTACCTGTGCAAAATGAAGA	3229
Db	3332	CCAGCTTAATTAATGAGATTTGGGCTGTGGAATGTGCACTACCTGTGCAAAATGAAGA	3391
QY	3230	AGCTGCTTGAGAGAGTGCAGCTACTCAAGCCCCGAGTGTAAATTGACGGTATGTTGGCAT	3289
Db	3392	AGCTGCTTGAGAGAGTGCAGCTACTCAAGCCCCGAGTGTAAATTGACGGTATGTTGGCAT	3451
QY	3290	CTTTGATGAAGATGACCG 3308	
Db	3452	CTTTGATGAAGATGACCG 3470	
RESULT 2	AY154460	3885 bp	mRNA linear PRI 20-FEB-2003
LOCUS	AY154460		
DEFINITION	Homo sapiens NALP5 (NALP5) mRNA, complete cds.		
ACCESSION	AY154460		
VERSION	AY154460.1	GI:28436363	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 3885)		
TITLE	Techopp,J., Martinson,F. and Burns,K.		
JOURNAL	NALP5: a novel protein family involved in inflammation		
MEDLINE	Nat. Rev. Mol. Cell Biol. 4 (2), 95-104 (2003)		
PUBMED	12563287		
REFERENCE	2 (bases 1 to 3885)		
AUTHORS	Martinson,F., Hofmann,K. and Techopp,J.		
TITLE	Direct Submision		
JOURNAL	Submitted (25-SEP-2002) Institute of Biochemistry, University of Lausanne, ch. des Boveresses 155, Epalinges, VD 1066, Switzerland		
FEATURES	Location/Qualifiers		
source	1..3885		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	1..3885		

[illegible]

QY 830 CTGCTCAAGATGACAAAGCTCTGCAAAAGCTGGGGCTGAGAGAGCGCTCCGTTCA 889  
Db 1106 CTGCTCTCAACATGACAAAGCTCTGCAAAAGCTGGGGCTGAGAGAGCGCTCCGTTCA 1165  
QY 890 CCTTCATACGAGTCTGCTGAGAGAGGTCCTGCTCCCTGAGTCTTCTGATCTGTACCG 949  
Db 1166 CCTTCATACGAGTCTGCTGAGAGAGGTCCTGCTCCCTGAGTCTTCTGATCTGTACCG 1225  
QY 950 TCAGAGAGTGGGGCAAGAGAGCTCAGTCAAGAGTGTGTCTCCCGGTTAACCTGTTAG 1009  
Db 1226 TCAGAGAGTGGGGCAAGAGAGCTCAGTCAAGAGTGTGTCTCCCGGTTAACCTGTTAG 1285  
QY 1010 TTAGAGGATCTCCGGGGAAAGAAATCACTTGCTCTTGAGCGCGGATGTGTAGC 1069  
Db 1286 TTAGAGGATCTCCGGGGAAAGAAATCACTTGCTCTTGAGCGCGGATGTGTAGC 1345  
QY 1070 ATCAGAGACACAGGGTGTGGTGCATATGAAACACCGTGAAGTCTCCACAGTCC 1129  
Db 1346 ATCAGAGACACAGGGTGTGGTGCATATGAAACACCGTGAAGTCTCCACAGTCC 1405  
QY 1130 AGGTGCGCGCGTGGGCTCTGTCATCTGGGGCCCTGAGCTGACGAGCGTGGGG 1189  
Db 1406 AGGTGCGCGCGTGGGCTCTGTCATCTGGGGCCCTGAGCTGACGAGCGTGGGG 1465  
QY 1190 AGAGGTCGCGCGCTTCAACCAAGCTCAAGAGCTGACGCGCTTTGTGTTCATC 1249  
Db 1466 AGAGGTCGCGCGCTTCAACCAAGCTCAAGAGCTGACGCGCTTTGTGTTCATC 1525  
QY 1250 AGCTCACCCCTGAGAGGCTGTGTCGGCGCTGTCTCATCTGAGAGAAAGTGTCTGA 1309  
Db 1526 AGCTCACCCCTGAGAGGCTGTGTCGGCGCTGTCTCATCTGAGAGAAAGTGTCTGA 1585  
QY 1310 AGCGCTTCTGCGATGAGCTGTGAGAGGAGTGTGAATAGAAAGTCACTGTTTGAAGT 1369  
Db 1586 AGCGCTTCTGCGATGAGCTGTGAGAGGAGTGTGAATAGAAAGTCACTGTTTGAAGT 1645  
QY 1370 ACGACCTCATGATTCAGAGACTCGGGAGTCTGAGCTCCGTCTCTGTTTCAATGAA 1429  
Db 1646 ACGACCTCATGATTCAGAGACTCGGGAGTCTGAGCTCCGTCTCTGTTTCAATGAA 1705  
QY 1430 TCTCTTCTCCAGACAGCACTGTGTAGAGTACTACACTTCTTCCACTCACTGTCCAG 1489  
Db 1706 TCTCTTCTCCAGACAGCACTGTGTAGAGTACTACACTTCTTCCACTCACTGTCCAG 1765  
QY 1490 ACTTCTGCGCGCTTGTACTACGTGTAGAGGGCTGGAATGAGGCAAGCTCTGACC 1549  
Db 1766 ACTTCTGCGCGCTTGTACTACGTGTAGAGGGCTGGAATGAGGCAAGCTCTGACC 1825  
QY 1550 CTCTGTACGTTGAGAGACAAAGAGGTCTATGAGCTTAAACAGGCAAGCTTCCATATC 1609  
Db 1826 CTCTGTACGTTGAGAGACAAAGAGGTCTATGAGCTTAAACAGGCAAGCTTCCATATC 1885  
QY 1610 ACTGCTTTGATGAGAGGTTTCTTGTGGCTCTGTAGAGAGAGTAAAGAGGCGAC 1669  
Db 1886 ACTGCTTTGATGAGAGGTTTCTTGTGGCTCTGTAGAGAGAGTAAAGAGGCGAC 1945  
QY 1670 TGAAGGTCCTGCTGGGCTGTCCCGTTCCCTGGGGGTGAGAGAGGTTTCTGACCTGG 1729  
Db 1946 TGAAGGTCCTGCTGGGCTGTCCCGTTCCCTGGGGGTGAGAGAGGTTTCTGACCTGG 2005  
QY 1730 TCTCTCTGTTGGGTCAGAGCTTAATGCAACCCAGAGAGACACCTTGAAGCGCTTCC 1789  
Db 2006 TCTCTCTGTTGGGTCAGAGCTTAATGCAACCCAGAGAGACACCTTGAAGCGCTTCC 2065  
QY 1790 ACTGCTTTTGGAGCTCAAGACAAAGATTGTGCTTGGCATTTAAACAGCTTCCAG 1849  
Db 2066 ACTGCTTTTGGAGCTCAAGACAAAGATTGTGCTTGGCATTTAAACAGCTTCCAG 2125  
QY 1850 AAGGTGCTTCCGATTAACAGAGACCTGGAATTGATAGCATTTCTTCCCTCCAGC 1909  
Db 2126 AAGGTGCTTCCGATTAACAGAGACCTGGAATTGATAGCATTTCTTCCCTCCAGC 2185  
QY 1910 ACTGTCCGTATTTGGGAAAAATTGGGTGATGTCAAGAGATCTTCCCAAGATGAGT 1969

Db 2186 ACTGTCCGTATTTGCGGAAAAATTGGGTGATGTCAAGGAGTCTTCCAAAGATGAGT 2245  
QY 1970 CCGCTGAGGCACTGTCTGTGTGTCCTCTATAGATGCGGATTAACACCTCTCATTAAGAGC 2029  
Db 2246 CCGCTGAGGCACTGTCTGTGTGTCCTCTATAGATGCGGATTAACACCTCTCATTAAGAGC 2305  
QY 2030 AGTGGGAAGATTTTGTGCTCCATGCTTGAGACCCACCAACCTGCGAGCTGAGCTG 2089  
Db 2306 AGTGGGAAGATTTTGTGCTCCATGCTTGAGACCCACCAACCTGCGAGCTGAGCTG 2365  
QY 2090 GCAGCAGATCTGACAGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGCATCCCA 2149  
Db 2366 GCAGCAGATCTGACAGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGCATCCCA 2425  
QY 2150 CCTGCAAGATPACACCTCTGATGTTTGAATGACAGATTAACCCCTGTGTGAGACAC 2209  
Db 2426 CCTGCAAGATPACACCTCTGATGTTTGAATGACAGATTAACCCCTGTGTGAGACAC 2485  
QY 2210 TCTGAGATGCTCATGGCCCAACCGTAACCTAAGATCCCTCACTTGGAGGCAACCA 2269  
Db 2486 TCTGAGATGCTCATGGCCCAACCGTAACCTAAGATCCCTCACTTGGAGGCAACCA 2545  
QY 2270 TGAAGAAAGAGATGTAAGATGCGGTGTGAAGCTTAAACACCCAAATGTTTGTGG 2329  
Db 2546 TGAAGAAAGAGATGTAAGATGCGGTGTGAAGCTTAAACACCCAAATGTTTGTGG 2605  
QY 2330 AGTCTTTGAGGCTGGAATTTGCTGTGTGATTTGACCCCATGCTGTACTGAAGATCCCA 2389  
Db 2606 AGTCTTTGAGGCTGGAATTTGCTGTGTGATTTGACCCCATGCTGTACTGAAGATCCCA 2665  
QY 2390 TCTTTACGACCTCCCAAGCGCTGAAATCTGAGCTGCGAGAAACAGAGTGAACAGC 2449  
Db 2666 TCTTTACGACCTCCCAAGCGCTGAAATCTGAGCTGCGAGAAACAGAGTGAACAGC 2725  
QY 2450 AGGAGTATATGCTCTGATGATGCTTTGAGAGTCTCCAGTGGCCCTGCAAGACTGA 2509  
Db 2726 AGGAGTATATGCTCTGATGATGCTTTGAGAGTCTCCAGTGGCCCTGCAAGAGCTGA 2785  
QY 2510 TACTGAGAGACTGTGGATGACACACCGGTTTCCAGAGCTGCGGCTCCAGCCCTGCA 2569  
Db 2786 TACTGAGAGACTGTGGATGACACACCGGTTTCCAGAGCTGCGGCTCCAGCCCTGCA 2845  
QY 2570 GCAACCGGAGCTTGAACACCTGTGCTATTCACCAACAGGCTGGGAAACGAGTGTAA 2629  
Db 2846 GCAACCGGAGCTTGAACACCTGTGCTATTCACCAACAGGCTGGGAAACGAGTGTAA 2905  
QY 2630 ATCTACTGTGTCATTCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATGCTGAATC 2689  
Db 2906 ATCTACTGTGTCATTCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATGCTGAATC 2965  
QY 2690 AGTGCACCTGGAACACCGGCTGGCTGTGGTTTCTTGAACCTTGAGGTTAACTCAT 2749  
Db 2966 AGTGCACCTGGAACACCGGCTGGCTGTGGTTTCTTGAACCTTGAGGTTAACTCAT 3025  
QY 2750 GGTGAGGCACTGAGCTTATGATGAACCTGTGTGAAGACAAATGSGGTGAAGCTTGT 2809  
Db 3026 GGTGAGGCACTGAGCTTATGATGAACCTGTGTGAAGACAAATGSGGTGAAGCTTGT 3085  
QY 2810 GCGAGTCAATGAGAAACCATTTGTCAATCTCCAGACCTGAGATTGATTAAGTCAATC 2869  
Db 3086 GCGAGTCAATGAGAAACCATTTGTCAATCTCCAGACCTGAGATTGATTAAGTCAATC 3145  
QY 2870 TCACCGCGGCTGTGTGAGAGTCTGTCTGTGTGATCTGAGAGACGACACCTGAAGA 2929  
Db 3146 TCACCGCGGCTGTGTGAGAGTCTGTCTGTGTGATCTGAGAGACGACACCTGAAGA 3205  
QY 2930 GCTGTGATCTACAGGACAAATGCTCTGGGTGAACGCTGTGCGCTGTGCGAGGAC 2989  
Db 3206 GCTGTGATCTACAGGACAAATGCTCTGGGTGAACGCTGTGCGCTGTGCGAGGAC 3265  
QY 2990 TGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGACGTGTGACTGACTTCTG 3049



Db 3266 TGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGATGTGACTGACTTCTG 3325  
Qy 3050 ATTTGCTGAGGCACTCTCTCTGCGCCCTTTCTTGCAACCGGATCTGACACTTAAAC 3109  
Db 3336 ATTTGCTGAGGCACTCTCTCTGCGCCCTTTCTTGCAACCGGATCTGACACTTAAAC 3385  
Qy 3110 TGTGTCAGAAATACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 3169  
Db 3386 TGTGTCAGAAATACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 3445  
Qy 3170 CCAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 3229  
Db 3446 CCAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 3505  
Qy 3320 AGCTCTGAGGAGAGTGCAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 3289  
Db 3506 AGCTCTGAGGAGAGTGCAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 3565  
Qy 3290 CTTTGTATGAGATGACCG 3308  
Db 3566 CTTTGTATGAGATGACCG 3584

RESULT 3  
AX704821 3926 bp DNA linear PAT 04-APR-2003  
DEFINITION Sequence 1 from Patent EP1285964.  
ACCESSION AX704821  
VERSION AX704821.1 GI:29561487  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1  
Welles, B., Lesel, M., Peters-Kottig, M. and Beckmann, G.  
AUTHORS Human mater proteins  
TITLE Patent: EP 1285964-A 1 26-FEB-2003;  
JOURNAL SCHERING AKTIBIOSELSCHAFT (DE)  
FEATURES  
source 1..3926  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 72.3%; Score 2917; DB 6; Length 3926;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 650 ACCAGGGAATGTTCTCTACAGCTCTTCTCTCCCGTTAGAGAGTGAAGGGAAGG 709  
Db 812 ACCAGGGAATGTTCTCTACAGCTCTTCTCTCCCGTTAGAGAGTGAAGGGAAGG 871  
Qy 710 AGAGCAGTGTACAGAGTTCATCTCAGAGAGTGGCCAGACTCCAGGCTCCGATGACG 769  
Db 872 AGAGCAGTGTACAGAGTTCATCTCAGAGAGTGGCCAGACTCCAGGCTCCGATGACG 931  
Qy 770 AGATCATGTCCCGACAGAAAGGCTGTTGTTATCATTTAGACGTTTGCATGACTGGCT 829  
Db 932 AGATCATGTCCCGACAGAAAGGCTGTTGTTATCATTTAGACGTTTGCATGACTGGCT 991  
Qy 830 CTGTCTCAAGAAATGACCAAGAGCTGCAAGAGCTGGGCTGAGAGAGAGCTCCGTTCA 889  
Db 992 CTGTCTCAAGAAATGACCAAGAGCTGCAAGAGCTGGGCTGAGAGAGAGCTCCGTTCA 1051  
Qy 890 CCTTCATAGCAGAGTGTCTGAGAGAGTCTGCTCCCTGAGTCTTCTGATCTGACCG 949  
Db 1052 CCTTCATAGCAGAGTGTCTGAGAGAGTCTGCTCCCTGAGTCTTCTGATCTGACCG 1111  
Qy 950 TCAGAGAGTGGGACAGAGAGCTCAAGTCAAGAGTGTCTTCCCGTTACCTGTTAG 1009  
Db 1112 TCAGAGAGTGGGACAGAGAGCTCAAGTCAAGAGTGTCTTCCCGTTACCTGTTAG 1171  
Qy 1010 TTAGAGAGATCTCCGAGGAGAAAGAAATCCATCTGCTCTTGAAGCGGAGTTGTGAGC 1065  
Db 1172 TTAGAGAGATCTCCGAGGAGAAAGAAATCCATCTGCTCTTGAAGCGGAGTTGTGAGC 1231  
Qy 1070 ATCAGAGAGCAAGAGGTTGGTGGATCATGAACAACGAGTGTCTGACCGAGTGC 1129  
Db 1232 ATCAGAGAGCAAGAGGTTGGTGGATCATGAACAACGAGTGTCTGACCGAGTGC 1291  
Qy 1130 AGGTCCCGCGGTGGGCTCTCTCATCTGCGTGGCCCTGACAGTGCAGAGCGTGGGG 1189  
Db 1292 AGGTCCCGCGGTGGGCTCTCTCATCTGCGTGGCCCTGACAGTGCAGAGCGTGGGG 1351  
Qy 1190 AGAGGTGGCCCGCTTCAACCAAGCTCAAGAGCTGACAGCGCTTTGTGTTTATC 1249  
Db 1352 AGAGGTGGCCCGCTTCAACCAAGCTCAAGAGCTGACAGCGCTTTGTGTTTATC 1411  
Qy 1250 AGCTACCCCTGAGAGCGTGGTCCGAGCTGCTCAATCTGAGAGAAAGTGTCTGCA 1309  
Db 1412 AGCTACCCCTGAGAGCGTGGTCCGAGCTGCTCAATCTGAGAGAAAGTGTCTGCA 1471  
Qy 1310 AGCGCTTCCGCTATGAGTGTGAGAGGAGTGTGAATGAGAGTCACTGTTTATGAGT 1369  
Db 1472 AGCGCTTCCGCTATGAGTGTGAGAGGAGTGTGAATGAGAGTCACTGTTTATGAGT 1531  
Qy 1370 ACGACTCATGTTCAAGAGTCTCGGAGAGTCTGAGCTCGTCTCTGTTTCAATGAACA 1429  
Db 1532 ACGACTCATGTTCAAGAGTCTCGGAGAGTCTGAGCTCGTCTCTGTTTCAATGAACA 1591  
Qy 1430 TCTCTTCCGAGACAGCCATGTAAGAGTCTACCTTCTTCACTTCACTGCTCCAG 1489  
Db 1592 TCTCTTCCGAGACAGCCATGTAAGAGTCTACCTTCTTCACTTCACTGCTCCAG 1651  
Qy 1490 ACTTCTGCGCGCTTGTACTAGTGTGAGAGGCTGGAATGAGAGCTCACTCTGCC 1549  
Db 1652 ACTTCTGCGCGCTTGTACTAGTGTGAGAGGCTGGAATGAGAGCTCACTCTGCC 1711  
Qy 1550 CTCTGTACGTTGAGAGCAAGAGTCTATGAGCTTAAACAGGCAAGCTTCAATATC 1609  
Db 1712 CTCTGTACGTTGAGAGCAAGAGTCTATGAGCTTAAACAGGCAAGCTTCAATATC 1771  
Qy 1610 ACTGCTTTGAGATGAGGTTCTTGTGTTGCTCTGAGAGAGAGCTTAAAGAGGCTAC 1669  
Db 1772 ACTGCTTTGAGATGAGGTTCTTGTGTTGCTCTGAGAGAGAGCTTAAAGAGGCTAC 1831  
Qy 1670 TGAAGTCTGCTGAGGCTGTCCGTTCCCTGAGGAGTGAAGCAAGCTTCTGACTGG 1729  
Db 1832 TGAAGTCTGCTGAGGCTGTCCGTTCCCTGAGGAGTGAAGCAAGCTTCTGACTGG 1891

Qy 1730 TCTCTGTGTGGGTGAGAGAGCTTAATGCCACACCCAGAGAGACCTCTGAGCGCTTCC 1789  
Db 1892 TCTCTGTGTGGGTGAGAGAGCTTAATGCCACACCCAGAGAGACCTCTGAGCGCTTCC 1951  
Qy 1790 ACTGTCTTTTGAAGCTCAAGACAAAGATTGTTGCTTGAGCATTAACAGCTTCCAG 1849  
Db 1952 ACTGTCTTTTGAAGCTCAAGACAAAGATTGTTGCTTGAGCATTAACAGCTTCCAG 2011  
Qy 1850 AAGTGTGCTTCCGATTAAACAGAACTTGATAGCATCTTCTGCTGCCAGC 1909  
Db 2012 AAGTGTGCTTCCGATTAAACAGAACTTGATAGCATCTTCTGCTGCCAGC 2071  
Qy 1910 ACTGTCTTTTGAAGCTCAAGAAATCGGTGATGTCAAGAGATCTTCCAGAGATAGT 1969  
Db 2072 ACTGTCTTTTGAAGCTCAAGAAATCGGTGATGTCAAGAGATCTTCCAGAGATAGT 2131  
Qy 1970 CCGCTGAGGCAATGCTGTGTGCTCTCTATAGATGCGGATTAAGACCTCATTTAGAGAGC 2029  
Db 2132 CCGCTGAGGCAATGCTGTGTGCTCTCTATAGATGCGGATTAAGACCTCATTTAGAGAGC 2191  
Qy 2030 AGTGGGAAGATTTTGGCTCCAGTGTGGACCCACCTGCGGAGCTGGACCTGG 2089  
Db 2192 AGTGGGAAGATTTTGGCTCCAGTGTGGACCCACCTGCGGAGCTGGACCTGG 2251  
Qy 2090 GCAGCAGCATCTGACAGAGCGGCGCATGAAGACCTGTGTGCAAGCTGAGGATCCCA 2149  
Db 2252 GCAGCAGCATCTGACAGAGCGGCGCATGAAGACCTGTGTGCAAGCTGAGGATCCCA 2311  
Qy 2150 CCTGCAGATACAGACCTGTGATTTAGAAATGCACAGATTACCTCTGTGTGAGCAGC 2209  
Db 2312 CCTGCAGATACAGACCTGTGATTTAGAAATGCACAGATTACCTCTGTGTGAGCAGC 2371  
Qy 2210 TCTGGAGATGCTGATGCGCAACCGTAACTTAAGTCCCTCAACTTGGGAGCAGCCAGC 2269  
Db 2372 TCTGGAGATGCTGATGCGCAACCGTAACTTAAGTCCCTCAACTTGGGAGCAGCCAGC 2431  
Qy 2270 TGAAGGAAGAGATGTAAGATGCGGTGAGAGCTTAAACACCCAAATGTTGTTGG 2329  
Db 2432 TGAAGGAAGAGATGTAAGATGCGGTGAGAGCTTAAACACCCAAATGTTGTTGG 2491  
Qy 2330 AGTCTTTGAGGCTGATGCTGTGATGATGACCCATGCTGTTAAGTAAATCTCCCAA 2389  
Db 2492 AGTCTTTGAGGCTGATGCTGTGATGATGACCCATGCTGTTAAGTAAATCTCCCAA 2551  
Qy 2390 TCTTACGACCTCCCGCAGCGCTGAATCTGAGCTGCGAGAAACAAAGGTGACAGC 2449  
Db 2552 TCTTACGACCTCCCGCAGCGCTGAATCTGAGCTGCGAGAAACAAAGGTGACAGC 2611  
Qy 2450 AGGAGTAATGCTCTCAGTATGCTTGAAGTCTCCAGTGCCTCTGAGAGAGCTGA 2509  
Db 2612 AGGAGTAATGCTCTCAGTATGCTTGAAGTCTCCAGTGCCTCTGAGAGAGCTGA 2671  
Qy 2510 TACTGAGAGACTGTGGCATCACAGCAGCGGTTGCGAAGTCTGAGCTTCCCTGCTCA 2569  
Db 2672 TACTGAGAGACTGTGGCATCACAGCAGCGGTTGCGAAGTCTGAGCTTCCCTGCTCA 2731  
Qy 2570 GCAACCCGAGCTTGACACACCTGCTTATCCAAACAGAGCTGGGAAACGAAGGTGAA 2629  
Db 2732 GCAACCCGAGCTTGACACACCTGCTTATCCAAACAGAGCTGGGAAACGAAGGTGAA 2791  
Qy 2630 ATCTACTGTGTGATCCTAGAGGCTTCCCACTGATGCTGAGAGGCTGATGCTGAATC 2689  
Db 2792 ATCTACTGTGTGATCCTAGAGGCTTCCCACTGATGCTGAGAGGCTGATGCTGAATC 2851  
Qy 2690 AGTGCACCTGAGACAGGCTGTGTGTTTCTTGACCTTGCGTTATGGGTAATCAT 2749  
Db 2852 AGTGCACCTGAGACAGGCTGTGTGTTTCTTGACCTTGCGTTATGGGTAATCAT 2911  
Qy 2750 GAGCTGAGCAGCTGAGCTTATGAGCAACCTGTGGAAGACAAATGGCGTGAAGCTTCTGT 2809  
Db 2912 GAGCTGAGCAGCTGAGCTTATGAGCAACCTGTGGAAGACAAATGGCGTGAAGCTTCTGT 2971  
Qy 2810 GCGAGTCAATGAGAACCATCTTGTCTCATCTTCAAGACCTGGAATGGTAAAGTCTATC 2869

Db 2972 GCGAGTCAATGAGAACCATCTTGTCTCATCTCCAGAGACTGAGATTGTAAGTGTATC 3031  
Qy 2870 TCACCGCGCGGTGTGTGAGAGTGTCTGCTGTGATCTGAGAGGAGACACCTGAGA 2929  
Db 3032 TCACCGCGCGGTGTGTGAGAGTGTCTGCTGTGATCTGAGAGGAGACACCTGAGA 3091  
Qy 2930 GCTGTGATCTCAAGCAATGCCCTGGGTGACAGGTGGGTTGTGCTGCTGTGCGAGGAC 2989  
Db 3092 GCTGTGATCTCAAGCAATGCCCTGGGTGACAGGTGGGTTGTGCTGCTGTGCGAGGAC 3151  
Qy 2990 TGAAGCAAAAGACAGTGTCTGACAGACTCGGTTGAAAGGACATGAGACTGACTTCTG 3049  
Db 3152 TGAAGCAAAAGACAGTGTCTGACAGACTCGGTTGAAAGGACATGAGACTGACTTCTG 3211  
Qy 3050 ATTGCTGTGAGGCACTCTCCCTTGCGCTTCCCTGCAACCGGCACTGACAGTAAAC 3109  
Db 3212 ATTGCTGTGAGGCACTCTCCCTTGCGCTTCCCTGCAACCGGCACTGACAGTAAAC 3271  
Qy 3110 TGTGTGAGAAATTAACCTTCAAGTCCCAAGAAATGATGAGTGTGTCGACCTTGTGTC 3169  
Db 3272 TGTGTGAGAAATTAACCTTCAAGTCCCAAGAAATGATGAGTGTGTCGACCTTGTGTC 3331  
Qy 3170 CCAGCTTAACCTTCAAGTGTGCTGTGAAATGCGAGTACCTGTGCAATTAAGA 3229  
Db 3332 CCAGCTTAACCTTCAAGTGTGCTGTGAAATGCGAGTACCTGTGCAATTAAGA 3391  
Qy 3230 AGCTGTGAGGAGAGTCAAGTCAAGCCCGAGTGTGTAATGAGGATGTTGGCAT 3289  
Db 3392 AGCTGTGAGGAGAGTCAAGTCAAGCCCGAGTGTGTAATGAGGATGTTGGCAT 3451  
Qy 3290 CTTTGTGATGAAGATGACCG 3308  
Db 3452 CTTTGTGATGAAGATGACCG 3470

RESULT 4  
AX704823 3830 bp DNA linear PAT 04-APR-2003  
LOCUS AX704823  
DEFINITION Sequence 3 from Patent EP1285964.  
ACCESSION AX704823  
VERSION AX704823.1 GI:29561488  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Human sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Weis, B., Lessl, M., Peters-Kotzig, M. and Beckmann, G.  
AUTHORS Human mater proteins  
TITLE Patent: EP 1285964-A 3 26-FEB-2003;  
JOURNAL SCHERING AKTIENSELSCHAFT (DE)  
FEATURES  
location/Qualifiers  
1..3830  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 70.9%; Score 2862; DB 6; Length 3830;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2962; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 345 AGAATTTCACAACCTATGGAACAAGAGTCCACAGCAGCAGAGACAGAAACAAG 404  
Db 450 AGAATTTCACAACCTATGGAACAAGAGTCCACAGCAGCAGAGACAGAAACAAG 509  
Qy 405 ACATGAGGTGACACATGAGGACTCAAGAGTCAAGTATGACCAATTCGTGTGAGAGA 464  
Db 510 ACATGAGGTGACACATGAGGACTCAAGAGTCAAGTATGACCAATTCGTGTGAGAGA 569  
Qy 465 GGATGTAGTGTGATGTTTGAACCACTGCTGCTGACTGGCCGGAATGCAACGTTGGC 524

Db 570 GGATGATACGTGATGTTTGAACACCTGCTGCTGATCGTCCGGAATGCAACGTGAGC 629  
Qy 525 TGGTGTCTTGAATTCAGACCGGTGGGGCTTCCGGCTCCGACGGTGTCTGACGGAA 584  
Db 630 TGGTGTCTTGAATTCAGACCGGTGGGGCTTCCGGCTCCGACGGTGTCTGACGGAA 689  
Qy 585 GTGAGGAATTTGGGAAATCGGCTCTAGCCAGAAAGATGCTGTGCTGGGGCCAAAGGTG 644  
Db 690 GTGAGGAATTTGGGAAATCGGCTCTAGCCAGAAAGATGCTGTGCTGGGGCCAAAGGTG 749  
Qy 645 ACTCTACAGGGAATGTTCTCTGACGCTCTTCTCTCTCCCGTATAGAGATGACGGAA 704  
Db 750 ACTCTACAGGGAATGTTCTCTGACGCTCTTCTCTCTCCCGTATAGAGATGACGGAA 809  
Qy 705 GAAAGAGACATGTCTACAGAGTTTCTCTCAGGAGTGGCCAGACTCCAGGCTCCGCT 764  
Db 810 GAAAGAGACATGTCTACAGAGTTTCTCTCAGGAGTGGCCAGACTCCAGGCTCCGCT 869  
Qy 765 GACGGAGATCATGTCCCGACAGAAAGGCTGTGTTTCACTTGAACGGTTTGATGACT 824  
Db 870 GACGGAGATCATGTCCCGACAGAAAGGCTGTGTTTCACTTGAACGGTTTGATGACT 929  
Qy 825 GGGCTCTGTCTCAACATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGACGCTCC 884  
Db 930 GGGCTCTGTCTCAACATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGACGCTCC 989  
Qy 885 GTTCAACCTCATACGCAAGTCTGTGAGAGAGTCTGTCTCTGAGTCTTCTGTATGCT 944  
Db 990 GTTCAACCTCATACGCAAGTCTGTGAGAGAGTCTGTCTCTGAGTCTTCTGTATGCT 1049  
Qy 945 CACCGTCAGAGAAGTGGGACAGAGAGTCAAGTCAAGAGTGTGTCTCCCGTTACT 1004  
Db 1050 CACCGTCAGAGAAGTGGGACAGAGAGTCAAGTCAAGAGTGTGTCTCCCGTTACT 1109  
Qy 1005 GTTATGTAAGGAATCTCCGGGGAACAAAGATCCACTTGTCTGTGACGGGATTTG 1064  
Db 1110 GTTATGTAAGGAATCTCCGGGGAACAAAGATCCACTTGTCTGTGACGGGATTTG 1169  
Qy 1065 TGAGCATCAGAAAGACAAAGGTTGGTGCATCATGAACAAACCTGAGCTCTGCACA 1124  
Db 1170 TGAGCATCAGAAAGACAAAGGTTGGTGCATCATGAACAAACCTGAGCTCTGCACA 1229  
Qy 1125 GTGCAAGGTGCGCGGTGGGCTCTCTACTGCGCGTGGCCGTCGACGTCAGAGAGTGT 1184  
Db 1230 GTGCAAGGTGCGCGGTGGGCTCTCTACTGCGCGTGGCCGTCGACGTCAGAGAGTGT 1289  
Qy 1185 GGGGAGAGAGCTGTGCGCCCTTCAACAAAGCTCAAGGCTGTGACGCGCTTTGTGT 1244  
Db 1290 GGGGAGAGAGCTGTGCGCCCTTCAACAAAGCTCAAGGCTGTGACGCGCTTTGTGT 1349  
Qy 1245 TCATGAGCTCACCCCTCGAGGCGTGTGCGGCTGTCTCAATCTGAGGAAAGATGT 1304  
Db 1350 TCATGAGCTCACCCCTCGAGGCGTGTGCGGCTGTCTCAATCTGAGGAAAGATGT 1409  
Qy 1305 CCTGAAGGCTTCTGCGCGTATGCTGTGAGAGAGTGTGAGATAGAGATGATGTTGA 1364  
Db 1410 CCTGAAGGCTTCTGCGCGTATGCTGTGAGAGAGTGTGAGATAGAGATGATGTTGA 1469  
Qy 1365 TGGTACGACATCATGTGTTCAAGAGCTCGGAGAGTCTGAGCTCCGCTGTCTGTCACT 1424  
Db 1470 TGGTACGACATCATGTGTTCAAGAGCTCGGAGAGTCTGAGCTCCGCTGTCTGTCACT 1529  
Qy 1425 GAAATCTCTTCTCCAGACAGCTGTGAGAGTACTACACTTCTTCAACTGAGTCT 1484  
Db 1530 GAAATCTCTTCTCCAGACAGCTGTGAGAGTACTACACTTCTTCAACTGAGTCT 1589  
Qy 1485 CCAGAGATCTGTGTGCGCGTGTACTAGTGTTAAGAGGCTGTGAGATGAGAGCTCT 1544  
Db 1590 CCAGAGATCTGTGTGCGCGTGTACTAGTGTTAAGAGGCTGTGAGATGAGAGCTCT 1649  
Qy 1545 CTGACCTCTGTAGTGAAGAGACAAAGAGTGCATGAGAGTCTTAACAGGAGCTTCA 1604  
Db 1650 CTGACCTCTGTAGTGAAGAGACAAAGAGTGCATGAGAGTCTTAACAGGAGCTTCA 1709

Qy 1605 TATCCACTGCTTNGATGAAGCCGTTCTGTGTTGGCTCGTGAAGCAAGACGTAAAGAG 1664  
Db 1710 TATCCACTGCTTNGATGAAGCCGTTCTGTGTTGGCTCGTGAAGCAAGACGTAAAGAG 1769  
Qy 1665 GCCACTGAGAGTCTGCTGAGGCTGTCCGCTTCCCTGGGGGGTGAAGAGAGCTTCTCA 1724  
Db 1770 GCCACTGAGAGTCTGCTGAGGCTGTCCGCTTCCCTGGGGGGTGAAGAGAGCTTCTCA 1829  
Qy 1725 CTGGTCTCTCTGTTGGGTCAAGACCTTAATGCCAACCCAGAGAGACACTTGAAGC 1784  
Db 1830 CTGGTCTCTCTGTTGGGTCAAGACCTTAATGCCAACCCAGAGAGACACTTGAAGC 1889  
Qy 1785 CTTCACATGCTTTTTCAGACTCAAGACAAAGAGTTTGTGCTTGGCTTGAACGCTT 1844  
Db 1890 CTTCACATGCTTTTTCAGACTCAAGACAAAGAGTTTGTGCTTGGCTTGAACGCTT 1949  
Qy 1845 CCAAGAGTGTGCTTCCGATTAACAGAACCTGACTTGAATGACTCTTCTGTGCT 1904  
Db 1950 CCAAGAGTGTGCTTCCGATTAACAGAACCTGACTTGAATGACTCTTCTGTGCT 2009  
Qy 1905 CCAGCACTGTCCGTAATTTGCGGAAATTCGGGTGATGTCAAGGATCTTCCAAAGAG 1964  
Db 2010 CCAGCACTGTCCGTAATTTGCGGAAATTCGGGTGATGTCAAGGATCTTCCAAAGAG 2069  
Qy 1965 TGAATCCGCTGAGGATGCTGTGTGTCTCTATGATGCGGATTAAGACCTTCAATTGA 2024  
Db 2070 TGAATCCGCTGAGGATGCTGTGTGTCTCTATGATGCGGATTAAGACCTTCAATTGA 2129  
Qy 2025 GAGAGAGTGGGAAATTTCTGCTCAATGCTTGGACCCACCCAGCTGGGAGCTTGA 2084  
Db 2130 GAGAGAGTGGGAAATTTCTGCTCAATGCTTGGACCCACCCAGCTGGGAGCTTGA 2189  
Qy 2085 CTTGGGAGAGCATCTCTGACAGAGCGGACATGAAGACCTGTGTGCAAGCTGAGCA 2144  
Db 2190 CTTGGGAGAGCATCTCTGACAGAGCGGACATGAAGACCTGTGTGCAAGCTGAGCA 2249  
Qy 2145 TCCACTGCAAGTACAGACCTGTGATTTGAAGATGACAGATTAACCTTGTGTGCA 2204  
Db 2250 TCCACTGCAAGTACAGACCTGTGATTTGAAGATGACAGATTAACCTTGTGTGCA 2309  
Qy 2205 GCACTCTGAGAGATCTGATGAGGCAACCGTAACCTTAAGATCCCTCAACTGGAGGAC 2264  
Db 2310 GCACTCTGAGAGATCTGATGAGGCAACCGTAACCTTAAGATCCCTCAACTGGAGGAC 2369  
Qy 2265 CCACCTGAAGAGAGATGTAAGAGTGGCTGTGAAGCTTTAAACACCCAAATGTTT 2324  
Db 2370 CCACCTGAAGAGAGATGTAAGAGTGGCTGTGAAGCTTTAAACACCCAAATGTTT 2429  
Qy 2325 GTTGAAGTCTTGAAGCTGATGCTGTGATGATGACCACTGTTAACCTGAAGATCTC 2384  
Db 2430 GTTGAAGTCTTGAAGCTGATGCTGTGATGATGACCACTGTGTAACTGAAGATCTC 2489  
Qy 2385 CCAATCTTTCAGACCTCCCGACCTGAATCTCTGAGCTGTGCAAGGAAACAAAGTAC 2444  
Db 2490 CCAATCTTTCAGACCTCCCGACCTGAATCTCTGAGCTGTGCAAGGAAACAAAGTAC 2549  
Qy 2445 AGACCAAGGAGTAAATGCTCTCAATGATGCTTGAAGATCTCCAGTGGCGCTGCAAG 2504  
Db 2550 AGACCAAGGAGTAAATGCTCTCAATGATGCTTGAAGATCTCCAGTGGCGCTGCAAG 2609  
Qy 2505 GCTAATCTGAGAGACTGTGAGCATCAAGCCAGGGTTGCCAGAGTCTGGCTTCAAGCCT 2564  
Db 2610 GCTAATCTGAGAGACTGTGAGCATCAAGCCAGGGTTGCCAGAGTCTGGCTTCAAGCCT 2669  
Qy 2565 GCTGAGCAACCGGAGCTTGAACACCTGTGCTTATCAACAAAGCTGTGGGAAACGAGG 2624  
Db 2670 GCTGAGCAACCGGAGCTTGAACACCTGTGCTTATCAACAAAGCTGTGGGAAACGAGG 2729  
Qy 2625 TGTAAATCTACTGTGATGATCATGAGGCTTCCCACTGTAGTCTGCAAGAGCTGATGCT 2684  
Db 2730 TGTAAATCTACTGTGATGATCATGAGGCTTCCCACTGTAGTCTGCAAGAGCTGATGCT 2789

Qy 2685 GAATCAGTCCACCTGACACAGGCTGCTGTTCTTCTGCACTTGGCTTATGGTAA 2744  
Db 2790 GAATCAGTCCACCTGACACAGGCTGCTGTTCTTCTGCACTTGGCTTATGGTAA 2849  
Qy 2745 CTCTAGGCTGACGCACTGAGCTTATGATGAACTCTGTGAAAGCAATGGCTGAGCT 2804  
Db 2850 CTCTAGGCTGACGCACTGAGCTTATGATGAACTCTGTGAAAGCAATGGCTGAGCT 2909  
Qy 2805 TCTGTGCAAGGTCATGAGAGAACCATCTGTATCTCCAGGACCTGGAGTGGTAAATG 2864  
Db 2910 TCTGTGCAAGGTCATGAGAGAACCATCTGTATCTCCAGGACCTGGAGTGGTAAATG 2969  
Qy 2865 TCATCTCACCGCCGCTGCTGTGAGAGTCTGTCTGTGTGATCTTCCAGAGACAGACCT 2924  
Db 2970 TCATCTCACCGCCGCTGCTGTGAGAGTCTGTCTGTGTGATCTTCCAGAGACAGACCT 3029  
Qy 2925 GAAAGAGCTGATCTCACGAGCAATGCCCTGGGTGACGCTGGGTTGCTGCTGCTGCGA 2984  
Db 3030 GAAAGAGCTGATCTCACGAGCAATGCCCTGGGTGACGCTGGGTTGCTGCTGCGA 3089  
Qy 2985 GGAAGCTGAAAGCAAGAGTCTGACGAGACTGGGTTGAAAGGACATGAGACTGAC 3044  
Db 3090 GGAAGCTGAAAGCAAGAGTCTGACGAGACTGGGTTGAAAGGACATGAGACTGAC 3149  
Qy 3045 TTCTGATTTCTGTGAGGCACTCTCCCTTGGCCCTTCTGCAACCGGCACTGACCACT 3104  
Db 3150 TTCTGATTTCTGTGAGGCACTCTCCCTTGGCCCTTCTGCAACCGGCACTGACCACT 3209  
Qy 3105 AAACCTGGTGCAGATTAATCTTCACTCCAAAGAAATGATGAAGCTGTGTGGCTTTGC 3164  
Db 3210 AAACCTGGTGCAGATTAATCTTCACTCCAAAGAAATGATGAAGCTGTGTGGCTTTGC 3269  
Qy 3165 CTGTCCCAAGCTCTAATCTTCAAGATTAATGGGCTGTGAAATGGAGTACCCCTGTGCAAT 3224  
Db 3270 CTGTCCCAAGCTCTAATCTTCAAGATTAATGGGCTGTGAAATGGAGTACCCCTGTGCAAT 3329  
Qy 3225 AAGGAGCTGCTGAGAGAGTGCAGCTACTCAAGCCCGAGTCTGTAAATTTGAACGTAATTG 3284  
Db 3330 AAGGAGCTGCTGAGAGAGTGCAGCTACTCAAGCCCGAGTCTGTAAATTTGAACGTAATTG 3389  
Qy 3285 GCATTCTTTTGTATGAAATGACCG 3308  
Db 3390 GCATTCTTTTGTATGAAATGACCG 3413

RESULT 5  
AX459881  
LOCUS AX459881 5859 bp DNA linear PAT 30-NOV-2002  
DEFINITION Sequence 24 from Patent WO240668.  
ACCESSION AX459881  
VERSION AX459881.1 GI:21725649  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Teschopp, J. and Martinon, F.  
TITLE Proteins and dna sequences underlying these proteins used for  
treating inflammations  
JOURNAL Patent: WO 0240668-A 24 23-MAY-2002;  
Apotech Research and Development Ltd. (CH)  
FEATURES  
SOURCE 1. 5859  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="NALP8/Pyl2---Pyl8 dna"

ORIGIN  
Query Match 69.5%; Score 2803; DB 6; Length 5859;  
Best Local Similarity .99.9%; Pred. No. 0;  
Matches 2903; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 290 AAATTTCAAGCTATGAAACAAGAGTGGCCACAGCAGAGAGACAGAAACAAGAAA 349  
Db 2287 AAATTTCAAGCTATGAAACAAGAGTGGCCACAGCAGAGAGACAGAAACAAGAAA 2446  
Qy 350 TTTCAAGCTATGAAACAAGAGTGGCCACAGCAGAGAGACAGAAACAAGAACATG 409  
Db 2447 TTTCAAGCTATGAAACAAGAGTGGCCACAGCAGAGAGACAGAAACAAGAACATG 2506  
Qy 410 GAGGTGACACATGGGACCTAACAAGTCAACGTATGACCAATTTGCTGAGAGAGAGATG 469  
Db 2507 GAGGTGACACATGGGACCTAACAAGTCAACGTATGACCAATTTGCTGAGAGAGAGATG 2566  
Qy 470 TACGTCGATGTTTGAAGAAACATGCTGTGACCTGGCCGAAATGCAACGTTGGCTG 529  
Db 2567 TACGTCGATGTTTGAAGAAACATGCTGTGACCTGGCCGAAATGCAACGTTGGCTG 2626  
Qy 530 CTTTGAATTCAGACCGGTGGGGCTTCCGGCTCGACAGTGTGTTCTGACCGAAAGTAC 589  
Db 2627 CTTTGAATTCAGACCGGTGGGGCTTCCGGCTCGACAGTGTGTTCTGACCGAAAGTAC 2686  
Qy 590 GAATTTGGGAAATGCGGCTCTAGCCAGAGATCGTGTGCGGCGCAAGTGGGACTCT 649  
Db 2687 GAATTTGGGAAATGCGGCTCTAGCCAGAGATCGTGTGCGGCGCAAGTGGGACTCT 2746  
Qy 650 ACCAGGAAATGTTCTCTTACGCTTCTTCTCCCTTGAAGAGATGACGAGAAAGAG 709  
Db 2747 ACCAGGAAATGTTCTCTTACGCTTCTTCTCCCTTGAAGAGATGACGAGAAAGAG 2806  
Qy 710 AGACAGTGTACAGAGTTCATCTCCAGAGAGTGGCCAGACTCCAGAGCTCGGTGACG 769  
Db 2807 AGACAGTGTACAGAGTTCATCTCCAGAGAGTGGCCAGACTCCAGAGCTCGGTGACG 2866  
Qy 770 AGATCATGTCCCGACCAAGAAAGCTGTGTTCAATTTGAACGTTTGAATGACCTGGCT 829  
Db 2867 AGATCATGTCCCGACCAAGAAAGCTGTGTTCAATTTGAACGTTTGAATGACCTGGCT 2926  
Qy 830 CTGTCTCAACAATGACACAAAGCTGTGAAAGAGTGGGCTGAGAACAGCTCTGTTCA 889  
Db 2927 CTGTCTCAACAATGACACAAAGCTGTGAAAGAGTGGGCTGAGAACAGCTCTGTTCA 2986  
Qy 890 CCTTCATACGAGTGTCTGTGAGAGAGTCTGTCTCTGAGTCTTCTGTATGCTCACCG 949  
Db 2987 CCTTCATACGAGTGTCTGTGAGAGAGTCTGTCTCTGAGTCTTCTGTATGCTCACCG 3046  
Qy 950 TCAGAGACGTGGGACAGAGAGTCAAGTCAAGAGTGTGTCTCCCGTTAAGCTGTAG 1009  
Db 3047 TCAGAGACGTGGGACAGAGAGTCAAGTCAAGAGTGTGTCTCCCGTTAAGCTGTAG 3106  
Qy 1010 TTAGAGGAAATCTCCGGGGAACAAGAAATGCACTTGTCTCTTGAAGCGGGAATTTGTAGC 1069  
Db 3107 TTAGAGGAAATCTCCGGGGAACAAGAAATGCACTTGTCTCTTGAAGCGGGAATTTGTAGC 3166  
Qy 1070 ATCAGAAAGACAAAGGTTGGTGTGATCATGAACAACGTGAGTGTGTGACAGTGGCC 1129  
Db 3167 ATCAGAAAGACAAAGGTTGGTGTGATCATGAACAACGTGAGTGTGTGACAGTGGCC 3226  
Qy 1130 AGGTGCGCGCGGTGGGCTCTCTATCTGCGGTGGCCGAGCTGACAGAGAGTGTGGGGG 1189  
Db 3227 AGGTGCGCGCGGTGGGCTCTCTATCTGCGGTGGCCGAGCTGACAGAGAGTGTGGGGG 3286  
Qy 1190 AGAGCGTGGCGCCCTTCAACCAACGCTCACAGGCTGACAGCGCTTTTGTGTTATC 1249  
Db 3287 AGAGCGTGGCGCCCTTCAACCAACGCTCACAGGCTGACAGCGCTTTTGTGTTATC 3346  
Qy 1250 AGCTACCCCTGAGGCGGTGTGCGGCTGTCTCAATCTGAGAGAAAGTGTCTGTA 1309  
Db 3347 AGCTACCCCTGAGGCGGTGTGCGGCTGTCTCAATCTGAGAGAAAGTGTCTGTA 3406  
Qy 1310 AGCGCTTCTGCGGTATGAGTGTGAGAGAGTGTGAGAAATAGAGTCAAGTGTGATG 1369  
Db 3407 AGCGCTTCTGCGGTATGAGTGTGAGAGAGTGTGAGAAATAGAGTCAAGTGTGATG 3466







Db	5507	AGCTTTGAGCGGTGATTTGCTGTGATTTGACCACTGCTGTTACTGAAGATCTCCAAA	5566
QY	2390	TCCTTTACGACCTTCCCCAGCGCTGAATCTGTAGCGCTG3CAGAGAAACAAGTGCACAGCC	2449
Db	5567	TCCTTTACGACCTTCCCCAGCGCTGAATCTGTAGCGCTG3CAGAGAAACAAGTGCACAGCC	5628
QY	2450	AGGGAGTAAATGCTCTCACTGATATGCTTGAAGTCTCCCAAGTGCCCTTGACGAAGCTGTA	2509
Db	5627	AGGGAGTAAATGCTCTCACTGATATGCTTGAAGTCTCCCAAGTGCCCTTGACGAAGCTGTA	5686
QY	2510	TACTGAGGACGTGTGGCATCAAGCAGCAGGGTTGCGAGAGTCTGCGCTCAGCGCTCGTCA	2569
Db	5687	TACTGAGGACGTGTGGCATCAAGCAGCAGGGTTGCGAGAGTCTGCGCTCAGCGCTCGTCA	5746
QY	2570	GCAACCGGAGCTTGAACAACCTGTGCTTATCCACACAGCGCTG9GAGAACAGATGTATA	2629
Db	5747	GCAACCGGAGCTTGAACAACCTGTGCTTATCCACACAGCGCTG9GAGAACAGATGTATA	5806
QY	2630	ATCTACTGTGTGATTCATAGAGGCTTCCCACTGTATGTCTGACAGGCTGTATGCTGAATC	2689
Db	5807	ATCTACTGTGTGATTCATAGAGGCTTCCCACTGTATGTCTGACAGGCTGTATGCTGAATC	5866
QY	2690	AGTGCACCTGGACACCGGTGGCTGTGGTTTTCTTGCACTTGCGCTTATGAGGTAACTCAT	2749
Db	5867	AGTGCACCTGGACACCGGTGGCTGTGGTTTTCTTGCACTTGCGCTTATGAGGTAACTCAT	5926
QY	2750	GCGTACGCACTGTAGCGCTTATAGCATGAACCTGTGTGAAGACATATGGCTGAAAGCTTCTGT	2809
Db	5927	GCGTACGCACTGTAGCGCTTATAGCATGAACCTGTGTGAAGACATATGGCTGAAAGCTTCTGT	5986
QY	2810	GCGAGGTCAATGAGAGAACATCTTGTCACTTCACAGACCTGTGATGGTAAAGTGTCAATC	2869
Db	5987	GCGAGGTCAATGAGAGAACATCTTGTCACTTCACAGACCTGTGATGGTAAAGTGTCAATC	6046
QY	2870	TCACCGCGCGCGTGTGATGAGTCTGTCTGTGTGATCTTCAGAGAGCAGACACTGGAAGA	2929
Db	6047	TCACCGCGCGCGTGTGATGAGTCTGTCTGTGTGATCTTCAGAGAGCAGACACTGGAAGA	6106
QY	2930	GCGTGTGATCTCAACGCAATATGCGCTGGGTGAACGAGTGGTGTCTGCGCTGTGCGAGGAC	2989
Db	6107	GCGTGTGATCTCAACGCAATATGCGCTGGGTGAACGAGTGGTGTCTGCGCTGTGCGAGGAC	6166
QY	2990	TGAACCAAAAGAACAGTGTCTTGACAGAGACTCGGGTTGAAGGCATGTGAACTGACTTCTG	3049
Db	6167	TGAACCAAAAGAACAGTGTCTTGACAGAGACTCGGGTTGAAGGCATGTGAACTGACTTCTG	6226
QY	3050	ATTGTGTGAGGCACTCTCTTGGCCCTTCTCTGAAACCGGACATCTGACAGCTTAACCC	3109
Db	6227	ATTGTGTGAGGCACTCTCTTGGCCCTTCTCTGAAACCGGACATCTGACAGCTTAACCC	6286
QY	3110	TGTCGCAAAATTAATCTTCAGTCCCAAGAAATGATGAAGCTGTGTCGCTTTGGCTGTCTC	3169
Db	6287	TGTCGCAAAATTAATCTTCAGTCCCAAGAAATGATGAAGCTGTGTCGCTTTGGCTGTCTC	6346
QY	3170	CCAGCTTAACCTTACAGATTAATTGG 3194	
Db	6347	CCAGCTTAACCTTACAGATTAATTGG 6371	
RESULT 7			
LOCUS	AX459891	6939 bp	DNA linear PAT 30-NOV-2002
DEFINITION	Sequence 34 from Patent WO0240668.		
ACCESSION	AX459891		
VERSION	AX459891.1	GI:21725654	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE			
AUTHORS	1		
TITLE	Tschoopp, J. and Martinon, F. Proteins and dna sequences underlying these proteins used for		

Query Match	69.5%	Score 2803	DB 6	Length 6939
Beet Local Similarity	99.9%	Pred. No. 0		
Matches 2903	Conservative	0	Mismatches 2	Indels 0
			Gaps 0	
ORIGIN	Creating inflammations Patent: WO 0240668-A 34 23-MAY-2002; Apotech Research and Development Ltd. (CH) Location/Qualifiers 1..6939 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" /note="NMRP13/Fy17.cdna"			
Query Match	69.5%	Score 2803	DB 6	Length 6939
Beet Local Similarity	99.9%	Pred. No. 0		
Matches 2903	Conservative	0	Mismatches 2	Indels 0
			Gaps 0	
ORIGIN	Creating inflammations Patent: WO 0240668-A 34 23-MAY-2002; Apotech Research and Development Ltd. (CH) Location/Qualifiers 1..6939 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" /note="NMRP13/Fy17.cdna"			

Dh 4307 AGGTGCCCCGGGGCTCTCATCTGTCGCTCCCTGACGTCGAGCAGCGTGTGGGG 4366  
Qy 1190 AGAGGTGCCCCCTTCAACCAAGCTCAAGGCTGCAAGCGCTTTGTGTTCATC 1249  
Dh 4367 AGAGGTGCCCCCTTCAACCAAGCTCAAGGCTGCAAGCGCTTTGTGTTCATC 4426  
Qy 1250 AGCTCACCCCTCGAGGCGTGTCCGGCGCTGTCTCAATCTGAGGAAAGAGTGTCTGA 1309  
Dh 4427 AGCTCACCCCTCGAGGCGTGTCCGGCGCTGTCTCAATCTGAGGAAAGAGTGTCTGA 4486  
Qy 1310 AGCGCTTGTCCGCTGTGCTGTGAGGAGTGTGGAATAGAAAGTCAAGTGTGTGATG 1369  
Dh 4487 AGCGCTTGTCCGCTGTGCTGTGAGGAGTGTGGAATAGAAAGTCAAGTGTGTGATG 4546  
Qy 1370 AGGACCTATGAGTTCAAGGACTCGGGGAGTCTGAGCTCGTGTCTGTTCACATGACA 1429  
Dh 4547 AGGACCTATGAGTTCAAGGACTCGGGGAGTCTGAGCTCGTGTCTGTTCACATGACA 4606  
Qy 1430 TCTCTTCTCCAGACAGCCACTGTGTGAGAGTACTACACCTTCTCCACTCAGTCTCAG 1489  
Dh 4607 TCTCTTCTCCAGACAGCCACTGTGTGAGAGTACTACACCTTCTCCACTCAGTCTCAG 4666  
Qy 1490 ACTTCTGTGCGGCTTGTACTACTGCTGTGAAGGCTGTGAAATGAGCTCCTGCTG 1549  
Dh 4667 ACTTCTGTGCGGCTTGTACTACTGCTGTGAAGGCTGTGAAATGAGCTCCTGCTG 4726  
Qy 1550 CTCTGTAGTTGAGAGCAAGAGGCTCATGAGCTTAAACAGCAGGCTTCCATATCC 1609  
Dh 4727 CTCTGTAGTTGAGAGCAAGAGGCTCATGAGCTTAAACAGCAGGCTTCCATATCC 4786  
Qy 1610 ACTGCTTTGAGTGAAGGCTTCTGTTTGGCTCTGTGAGGAGAACGTTAAGAGGCCAC 1669  
Dh 4787 ACTGCTTTGAGTGAAGGCTTCTGTTTGGCTCTGTGAGGAGAACGTTAAGAGGCCAC 4846  
Qy 1670 TGAAGTCTGTGCGGCTGTCCGTTCCCTGTGGGGGTGAAGCAAGCTTCTGCACTGG 1729  
Dh 4847 TGAAGTCTGTGCGGCTGTCCGTTCCCTGTGGGGGTGAAGCAAGCTTCTGCACTGG 4906  
Qy 1730 TCTCTGTGTGGGTGAGAGCTTATGTCACACCCGAGGAGACCCCTGAGCGCTTCC 1789  
Dh 4907 TCTCTGTGTGGGTGAGAGCTTATGTCACACCCGAGGAGACCCCTGAGCGCTTCC 4966  
Qy 1790 ACTGCTTTTGAAGCTCAAGACAAAGATTTGTGCTGTGCAATTAAACAGCTTCCAG 1849  
Dh 4967 ACTGCTTTTGAAGCTCAAGACAAAGATTTGTGCTGTGCAATTAAACAGCTTCCAG 5026  
Qy 1850 AAGTGTGCTTCCATTAACCAAGACCTGACCTTGATGCACTTCTTCTGCTCCAGC 1909  
Dh 5027 AAGTGTGCTTCCATTAACCAAGACCTGACCTTGATGCACTTCTTCTGCTCCAGC 5086  
Qy 1910 ACTGTCGATTTTGGGAAATTTGGGGTGAATGTCAAGAGGATCTTCCCAAGATGAGT 1969  
Dh 5087 ACTGTCGATTTTGGGAAATTTGGGGTGAATGTCAAGAGGATCTTCCCAAGATGAGT 5146  
Qy 1970 CCGCTGAGGCAATGCTGTGTGCTCCCTCATATGATGCGGGATTAAGACCTCTCATTTGAGAGC 2029  
Dh 5147 CCGCTGAGGCAATGCTGTGTGCTCCCTCATATGATGCGGGATTAAGACCTCTCATTTGAGAGC 5206  
Qy 2030 AGTGGAGAAATTTTGTCTTCATGCTTGGACACCCACACCTGTGCGGACGCTGACCTGG 2089  
Dh 5207 AGTGGAGAAATTTTGTCTTCATGCTTGGACACCCACACCTGTGCGGACGCTGACCTGG 5266  
Qy 2090 GCAAGAGATCTGACAGAGCGGGGCAATGAAGACCTGTGTGCGCAAGTGTGAGGATCCCA 2149  
Dh 5267 GCAAGAGATCTGACAGAGCGGGGCAATGAAGACCTGTGTGCGCAAGTGTGAGGATCCCA 5326  
Qy 2150 CCTGCAAGATACAGACCTGTATGTTAGAAATGACAGATTAACCTGTGTGACGACCC 2209  
Dh 5327 CCTGCAAGATACAGACCTGTATGTTAGAAATGACAGATTAACCTGTGTGTGACGACCC 5386  
Qy 2210 TCTGAGAAATGTCATGCGCAACCGTAACTTAAGATCTCTCACTTGTGGAGGCAACCCAC 2269  
Dh 5387 TCTGAGAAATGTCATGCGCAACCGTAACTTAAGATCTCTCACTTGTGGAGGCAACCCAC 5446

Qy 2270 TGAAGAAAGAGATGTAAGATGCGTGTGAAGCTTAAACACCCAAATGTTTGTGG 2329  
Dh 5447 TGAAGAAAGAGATGTAAGATGCGTGTGAAGCTTAAACACCCAAATGTTTGTGG 5506  
Qy 2330 AGTCTTGAAGGTGGAATTTGTGTGATTTGACCCATGCTGTATCTGAAGATCTCCAAA 2389  
Dh 5507 AGTCTTGAAGGTGGAATTTGTGTGATTTGACCCATGCTGTATCTGAAGATCTCCAAA 5566  
Qy 2390 TCTCTTGAAGCTTCCCGAGCTGAAATCTGTGAGGCTGTGGAGGAAACAAGTGAACAACC 2449  
Dh 5567 TCTCTTGAAGCTTCCCGAGCTGAAATCTGTGAGGCTGTGGAGGAAACAAGTGAACAACC 5626  
Qy 2450 AGGAGTAATGCTCTCATGATATGCTTGAAGAGTCTCCAGTGGGCCCTGACAGAGCTGA 2509  
Dh 5627 AGGAGTAATGCTCTCATGATATGCTTGAAGAGTCTCCAGTGGGCCCTGACAGAGCTGA 5686  
Qy 2510 TACTGAGAGCTGTGGATCAACAGCCAGGGGTTCAGAGTCTTGAGCTTCAAGCCTGTGCA 2569  
Dh 5687 TACTGAGAGCTGTGGATCAACAGCCAGGGGTTCAGAGTCTTGAGCTTCAAGCCTGTGCA 5746  
Qy 2570 GCAACCGGAGCTTGAACAACCTGTGCTTATCCAAACAACCTGTGGGAAAGAAAGTTAA 2629  
Dh 5747 GCAACCGGAGCTTGAACAACCTGTGCTTATCCAAACAACCTGTGGGAAAGAAAGTTAA 5806  
Qy 2630 ATTACTGTGTGCATCCATGAGGCTTCCCACTGTATCTGACAGAGCTGATGCTGAATC 2689  
Dh 5807 ATTACTGTGTGCATCCATGAGGCTTCCCACTGTATCTGACAGAGCTGATGCTGAATC 5866  
Qy 2690 AGTGCCACTGTGACACGGGCTGTGTGTTTCTTGACATTTGCGCTTATGGGTAACTCAT 2749  
Dh 5867 AGTGCCACTGTGACACGGGCTGTGTGTTTCTTGACATTTGCGCTTATGGGTAACTCAT 5926  
Qy 2750 GGTCTGAAGCACTTGAAGCTTATAGATGAACCTGTGGAAGCAATGTGGTAAAGCTTCTGT 2809  
Dh 5927 GGTCTGAAGCACTTGAAGCTTATAGATGAACCTGTGGAAGCAATGTGGTAAAGCTTCTGT 5986  
Qy 2810 GCGAGTCTATGAGAGAACATCTTGTATCTCCAGACCTGTGAGATTTGTTAAAGTGTATC 2869  
Dh 5987 GCGAGTCTATGAGAGAACATCTTGTATCTCCAGACCTGTGAGATTTGTTAAAGTGTATC 6046  
Qy 2870 TCACCGCGGCTGTGTGAGAGTGTCTGTGTGATCTTGAGAGAGACACCTGAAAGA 2929  
Dh 6047 TCACCGCGGCTGTGTGAGAGTGTCTGTGTGATCTTGAGAGAGACACCTGAAAGA 6106  
Qy 2930 GCTGTGATCTCAACGGAATGCTGTGGTGAACGCTGTGGGTGTGCTGCCCTGTGAGAGGAC 2989  
Dh 6107 GCTGTGATCTCAACGGAATGCTGTGGTGAACGCTGTGGGTGTGCTGCCCTGTGAGAGGAC 6166  
Qy 2990 TGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCAATGTGACTGACTTCTG 3049  
Dh 6167 TGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGCAATGTGACTGACTTCTG 6226  
Qy 3050 ATTGTGTGAGGCACTCTCTTGGCCCTTCTGTCAACCGGCATCTGACAGTCTAAACC 3109  
Dh 6227 ATTGTGTGAGGCACTCTCTTGGCCCTTCTGTCAACCGGCATCTGACAGTCTAAACC 6286  
Qy 3110 TGTGTGCAATTAATTTCAATGCTCCAAAGGAATGATGAAGCTGTGTGCGCTTGTGCTGTC 3169  
Dh 6287 TGTGTGCAATTAATTTCAATGCTCCAAAGGAATGATGAAGCTGTGTGCGCTTGTGCTGTC 6346  
Qy 3170 CCAGTCTAACTTACAGATTAATTGG 3194  
Dh 6347 CCAGTCTAACTTACAGATTAATTGG 6371

RESULT 8  
AY054986 3885 bp mRNA linear PRI 02-APR-2002  
LOCUS Homo sapiens maternal-antigen-that-embryos-require protein (MATER)  
DEFINITION mRNA, complete cds.  
ACCESSION AY054986  
VERSION AY054986.1 GI:19882272



KEYWORDS	SOURCE
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE	1 (bases 1 to 3885)
JOURNAL	Tong, Z.-B., Bondy, C.A., Zhou, J. and Nelson, L.M.
MEDLINE	A human homologue of mouse Mater, a maternal effect gene essential for early embryonic development
PUBMED	Hum. Reprod. 17 (4), 903-911 (2002)
REFERENCE	21922687
AUTHORS	11925379
TITLE	2 (bases 1 to 3885)
JOURNAL	Tong, Z.-B., Bondy, C.A. and Nelson, L.M.
FEATURES	Direct Submission
	Submitted (30-AUG-2001) Developmental Endocrinology Branch, NICHD, NIH, 10 Center Drive, Bethesda, MD 20892/1862, USA
	Location/Qualifiers
	1..3885

gene	1. .3885
	/gene="WATER"
CDS	1. .3603

/note="oocyte-specific protein; similar to Mus musculus Mater encoded by GenBank Accession Numbers AF074018 and AF143559-AF143573"

```
/product="maternal-antigen-that-embryos-require protein
/protein_id="AAL15549.1"
/db_xref="GI:19882273"
```

SSQPCIMBGDKSLTFSSYGLQWCLYELDKKEFPOTFEKLLKKKSESTTCSIPOFT  
NANVECLALLHETTYGASLAWATSIISIFENNMLRTLSEKARDMDKRGHPEDPEATIS  
QGPSKEKVPGISQAVQODSATAAETKEQEISQAMEQEGATAAETBEQEISQAMEQ

FRPRVTVLHGKSGIGKSALARRITLCAOGGLYQGMFSEVFPLPVREMQRKKESSY  
FISREWPDQAPVTEIMSRPERLLFIIDGFDLGSVLNDTKLCKDWAEQPPFTIIT  
SLRKVLLPESFLIVTRDVCTEKLKSEVSPRYLLVRGISGEQRHLLLERGIGL

QUTPRGVVRCLNLEBRVVLKRFCRMAVEGVWNRKSVFDGDDLMVOGLGESELRBAL  
MNLILPDSHCBEYTFTHLSLODFCALYYVLBLEIBALCPLYVEKTRSMELL  
GFHHSLMKRELFGLVSEDEVRRPLEVLGCPVPLGVKQGLHWSLLGOONATTT

KGIFPRDESABACPVVPLMMDKTLIEBQWEDFCMGLGTHPHLRQDLGSSILTEE  
KTLCAKLRHPTCKIQTLMFRNAQITPGVQHLMRIVMANRNLRSLNGLGTHLKEED  
ACEALKHPKCLLESRLDCCGLTHACYLAKISOILTTPSLKSLSLAGNKVTDQGVV

RSMRLPHCSQRLMLNQCHLDTAGCGSLATALMGNSWLTHLSLSMNPVEDNGVYTL  
VMREPSCHLDQLELVKCHLTAAACESLCSVRSRHLKSLDLTDNALGDGVAALG  
LKQKNSVLTRLGLKACGLTSDCCEALSTALSCNRHLTSLNLVQNNFSPKGMKLCCL

ORIGIN	Query Match	67.2%	Score 2713	DB 9	Length 3885
--------	-------------	-------	------------	------	-------------

Matches 3013; Conservative 0; Mismatches 6; Indels 0; Gaps

566 A A A T T T C A C A A G C T A T G G A A C A A G A A G T G C C A C A G C A G C A G A G A C A G A A G A C A A G A A A 6  
350 T T T C A C A A G C T A T G G A A C A A G A A G T G C C A C A G C A G C A G A G A C A A G A A C A A G C A T G 4

626 TTTCACCACTATGSAACAGAAAGGTGCCACAGCAGCAGAGACAGAAACAAAGACATG 6  
410 GAGGTGACATGGGACTACAAAGCTCACCGTATGACCAATTCCGTGAGAGAGAGATG 4

686 GAGGTGACATGSGACTACAAGAGTCACCGTATGATGACCAATTCCGCTGAGGAGGAGGATG 74

Oy 470 TACGTCGAGTTTGGAAAAACACTGCTGTGA CTGCCCGAATGCAAACGTTGGCTGGTG 529  
| | | | | | | | | | | | | | | | | | | | | |  
Db 746 TACGTCGAGTTTGGAAAAACACTGCTGTGA CTGCCCGAATGCAAACGTTGGCTGGTG 805

Qy 530 CTTTGTGATTCAGACCGGTGGGGCTTCCGGCTCGACCGGTTCGTCACCGGAAAGTCAG 589

Db 806 CTTTGTGATTCAGACCGGTGGGGCTTCCGGCTCGACCGGTTCGTCACCGGAAAGTCAG 865

Qy 590 GAAATGGGAAATCGGCTTAGCCGAGGAGTGTGCTGTGGGCGCAAGGTGAACTCT 649  
|||  
Db 866 GAAATGGGAAATCGGCTTAGCCGAGGAGTGTGCTGTGGGCGCAAGGTGAACTCT 925

**DY**      650 ACCGGGAAATGTTCTCCTACGTCTTCTCCCGTAGAGAGTCAGCGGAAGAAGG 709  
|||||  
|||  
**Db**      926 ACCAGGGAATGTTCTCCTACGTCTTCTCCTCCCGTAGAGAGTCAGCGGAAGAAGG 985

QY 710 AGACGAGTGCACAGATTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGGTGAACGG 769  
 |||||  
 Db 986 AGACGAGTGCACAGATTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGGTGAACGG 1044

```

QY      770 AGATCATGTCGCCGACGAGAAAGCGTGTTCATCATTTGACGGTTTCGATGACCTGGGCT 829
          |||||
Db      1046 AGATCATGTCGCCGACGAGAAAGCGTGTTCATCATTTGACGGTTTCGATGACCTGGGCT 1100

```

Qy 830 CTGTCCTCAACAATGACACAAGCTCTGAAAGACTGGGGCTGAGAAAGCAGCCTCCGTTCA 889

Db 1106 CTGTCCTCAACAATGACACAAGCTCTGAAAGACTGGGGCTGAGAAAGCAGCCTCCGTTCA 1166

QY	890	CCCTCATAGCAGCTGCTGAGGAAAGTCTGCTCCCTGAGTCTTCTGATCTGACCG	949
Db	1166	CCCTCATAGCAGCTGCTGAGGAAAGTCTGCTCCCTGAGTCTTCTGATCTGACCG	1222

QY 950 TCAGAAGCGTGGGCACAGAGAAGCTCAAGTCAGAGGTCGTCTCCCCGTTACCTGTTAG 100

Db 1226 TCAGAAGCGTGGGCACAGAGAAGCTCAAGTCAGAGGTCGTCTCCCCGTTACCTGTTAG 128

Qy 1010 TTAAAGAACTCCGGGGAACAAGAATCACTTGCTCTTGAGCGGGGAAATTGTAAGC 106  
Db 1286 TTAAAGAACTCCGGGGAACAAGAATCACTTGCTCTTGAGCGGGGAAATTGTAAGC 134

Qy 1070 ATCAGAAGCACAAAGGTTCCGTGGATCATGAAACAACCGTAGCTCTCGAACAGTGC 1122

Db 1346 ATCAGAAGCACAAAGGTTCCGTGGATCATCAACAACCGTAGCTCTCGAACAGTGC 1408

1130 AGGAGCCCGCGTGGGCTCTCTCAACTGCGTGGCCCTGCAGCTGCAGGACGTTGGGGG 118  
 1406 AGGAGCCCGCGTGGGCTCTCTCAACTGCGTGGCCCTGCAGCTGCAGGACGTTGGGGG 146

1190 AGAGCGTGGCCCCCTTCAACCAACGCTACAGGCTGCACGCCGTTTGTTTCATC 124  
 1466 AGAGCGTGGCCCCCTTCAACCAACGCTACAGGCTGCACGCCGTTTGTTTCATC 1522

Qy 1250 AGCTCACCCCTCGAGGCGTGGTCCGGCGCTCTCAATCTGGAGGAAAGAGTTGTCTCTGA 130

Db 1526 AGCTCACCCCTCGAGGCGTGGTCCGGCGCTCTCAATCTGGAGGAAAGAGTTGTCTCTGA 158

QY 1310 AGCGCTTCGCCGATGCGCTGAGGGAATGTGGAATAGGAAGTCAGTGTTCGATGCTG 136

DB 1586 AGCGCTTCGCCGATGCGCTGAGGGAATGTGGAATAGGAAGTCAGTGTTCGATGCTG 164

1370 ACGACTCATGGTCAAGCACTCGGGAGCTGAGTCCGCTCTGTTTCAATGACA 1422  
 1646 ACGACTCATGGTCAAGCACTCGGGAGCTGAGTCCGCTCTGTTTCAATGACA 1700

Qy 1430 TCCTTCTCCAGACGCCACTGTGAGGAGTACTAACCTTCTTCCACTCAGTCTCCAGG 148  
Db 1706 TCCTTCTCCAGACGCCACTGTGAGGAGTACTAACCTTCTTCCACTCAGTCTCCAGG 176

1490 ACTCTGTCGCCCTGTGTAACGTTTGAAGGCGCTGGAATCGAGCCACTCTCTGCC 154  
 |||||  
 1766 ACTTCTGTGCGCCTTGTACTACGTGTTAGAGGCGCTGGAATCGAGCCACTCTCTGCC 182

1550 CTCTGACCTGGAGAGACAAAGAGGTCCATGGAGCTTAAACAGGCAGGCTTCCATATCC 1600

|||||  
Db 1826 CTTCTGATCTTGAGAAAGCAAGAGCTCATGAGCTTAAACAGGCAAGCTTCCATATCC 1885  
|||  
Qy 1610 ACTGCTTTGATGATGAGGCTTTCTTTGTTGCTCTGTGAGGAGAGAGCTAAGAGGACAC 1669  
|||  
Db 1886 ACTGCTTTGATGATGAGGCTTTCTTTGTTGCTCTGTGAGGAGAGAGCTAAGAGGACAC 1945  
|||  
Qy 1670 TGAAGGCTCTGCTGAGGCTGTCCCGTTCCCTGAGGAGGATGAGAGAGGCTTCTGACCTGG 1729  
|||  
Db 1946 TGAAGGCTCTGCTGAGGCTGTCCCGTTCCCGTTGAGGAGGATGAGAGAGGCTTCTGACCTGG 2005  
|||  
Qy 1730 TCTCTGTTGAGGCTGAGAGGCTTAAATGACCAACCCAGAGAGACACCTTGAAGGCTTCC 1789  
|||  
Db 2006 TCTCTGTTGAGGCTGAGAGGCTTAAATGACCAACCCAGAGAGACACCTTGAAGGCTTCC 2065  
|||  
Qy 1790 ACTGCTTTTGAAGACTCAAGACAAAGATTGTTGCTTGGCTTAAACAGCTTCCAG 1849  
|||  
Db 2066 ACTGCTTTTGAAGACTCAAGACAAAGATTGTTGCTTGGCTTAAACAGCTTCCAG 2125  
|||  
Qy 1850 AAGTGGCTTCCGATTAACAGAGCTGAGCTTGAATGACATCTTCTTGGCTCCAGC 1909  
|||  
Db 2126 AAGTGGCTTCCGATTAACAGAGCTGAGCTTGAATGACATCTTCTTGGCTCCAGC 2185  
|||  
Qy 1910 ACTGCTCGTATTTGGGAAATTCGGGTGATGTCAAAGGATCTTCCCAAGATGAGT 1969  
|||  
Db 2186 ACTGCTCGTATTTGGGAAATTCGGGTGATGTCAAAGGATCTTCCCAAGATGAGT 2245  
|||  
Qy 1970 CCGCTGAGGAGATGCTGTGCTCTCTATGATGACGGGATTAAGACCTTCAATGAGAGC 2029  
|||  
Db 2246 CCGCTGAGGAGATGCTGTGCTCTCTATGATGACGGGATTAAGACCTTCAATGAGAGC 2305  
|||  
Qy 2030 AGTGGAGAGATTTTGTCTCCATGCTTGGCAACCCACACCTGAGGAGCTGAGCTGG 2089  
|||  
Db 2306 AGTGGAGAGATTTTGTCTCCATGCTTGGCAACCCACACCTGAGGAGCTGAGCTGG 2365  
|||  
Qy 2090 GCAGCAGATCTCTGACAGAGCGGACCATGAAGACCTGTGTGCAAGCTGAGGATCCCA 2149  
|||  
Db 2366 GCAGCAGATCTCTGACAGAGCGGACCATGAAGACCTGTGTGCAAGCTGAGGATCCCA 2425  
|||  
Qy 2150 CCTGCAATACAGACCTCTGATGTTTAAAGATGACCAATTAACCTCTGTGTGAGAGC 2209  
|||  
Db 2426 CCTGCAATACAGACCTCTGATGTTTAAAGATGACCAATTAACCTCTGTGTGAGAGC 2485  
|||  
Qy 2210 TCTGAGATGATGCTGATGAGCAACCGTAACTTAAGTCCCTCAACTTGGAGGAGACCCAC 2269  
|||  
Db 2486 TCTGAGATGATGCTGATGAGCAACCGTAACTTAAGTCCCTCAACTTGGAGGAGACCCAC 2545  
|||  
Qy 2270 TGAAGGAGAGATGATGAGATGAGGCTGTAAGCTTAAACACCCAAATGTTTGTGG 2329  
|||  
Db 2546 TGAAGGAGAGATGATGAGATGAGGCTGTAAGCTTAAACACCCAAATGTTTGTGG 2605  
|||  
Qy 2330 AGTCTTTGAGGCTGATGCTGTGATGAGCAACCTGCTTAACTTGAAGATCTCCCAA 2389  
|||  
Db 2606 AGTCTTTGAGGCTGATGCTGTGATGAGCAACCTGCTTAACTTGAAGATCTCCCAA 2665  
|||  
Qy 2390 TCTCTACGACCTCCCCAGCCTGAATCTCTGAGGCTGAGAGAGAGAGAGAGAGAGAG 2449  
|||  
Db 2666 TCTCTACGACCTCCCCAGCCTGAATCTCTGAGGCTGAGAGAGAGAGAGAGAGAGAG 2725  
|||  
Qy 2450 AGGAGATATGCTCTGATGATGCTTGAAGTCTCCAGTGCCTTGCAGAGAGCTGA 2509  
|||  
Db 2726 AGGAGATATGCTCTGATGATGCTTGAAGTCTCCAGTGCCTTGCAGAGAGCTGA 2785  
|||  
Qy 2510 TACTGAGAGATCTGTGAGATCAACAGCAGGAGTGGCAAGATGAGCTCAAGCTTCCGCA 2569  
|||  
Db 2786 TACTGAGAGATCTGTGAGATCAACAGCAGGAGTGGCAAGATGAGCTCAAGCTTCCGCA 2845  
|||  
Qy 2570 GCAACCGGAGCTTGACACACCTGTGCTATCCAAACAGAGCTTGGAGAGAGAGAGTGA 2629  
|||  
Db 2846 GCAACCGGAGCTTGACACACCTGTGCTATCCAAACAGAGCTTGGAGAGAGAGAGTGA 2905  
|||  
Qy 2630 ATCTACTGTGCTGATCATGAGGCTTCCCACTGATGCTGCAAGAGGCTGATGCTGAATC 2689  
|||

Db 2906 ATCTACTGTGCTGATCATGAGGCTTCCCACTGTGATGCTGCAAGAGCTGATGATC 2965  
|||  
Qy 2690 AGTGCACCTGGAACAGGCTGAGCTGTGTTTCTTGGACCTTGAGCTTATGGATTAATCAT 2749  
|||  
Db 2966 AGTGCACCTGGAACAGGCTGAGCTGTGTTTCTTGGACCTTGAGCTTATGGATTAATCAT 3025  
|||  
Qy 2750 GGTGAGCAGCTTGAAGCTTATGATGAACCTGTGGAAGACAAATGAGGATGAGCTTGT 2809  
|||  
Db 3026 GGTGAGCAGCTTGAAGCTTATGATGAACCTGTGGAAGACAAATGAGGATGAGCTTGT 3085  
|||  
Qy 2810 GCGAGGTATGAGAGAACATCTTGTGATCTCCAGAGACTGAGAGATGATGAAGTGTATC 2869  
|||  
Db 3086 GCGAGGTATGAGAGAACATCTTGTGATCTCCAGAGACTGAGAGATGATGAAGTGTATC 3145  
|||  
Qy 2870 TCACGCGCGGTGTGATGAGAGTCTGTCTGTGATCTGAGAGAGAGACACCTGAAGA 2929  
|||  
Db 3146 TCACGCGCGGTGTGATGAGAGTCTGTCTGTGATCTGAGAGAGAGACACCTGAAGA 3205  
|||  
Qy 2930 GCTGATCTTCAACGAGCAATGCCCTGGGTGAACGATGGAGTGTGCTGCGCTGCGAGGAC 2989  
|||  
Db 3206 GCTGATCTTCAACGAGCAATGCCCTGGGTGAACGATGGAGTGTGCTGCGCTGCGAGGAC 3265  
|||  
Qy 2990 TGAAGCAAAAGACAGTGTCTGACAGAGACTCGGCTTGAAGGACATGTGAGACTGACTTG 3049  
|||  
Db 3266 TGAAGCAAAAGACAGTGTCTGACAGAGACTCGGCTTGAAGGACATGTGAGACTGACTTG 3325  
|||  
Qy 3050 ATTGCTGTGAGGACATCTCTCTGAGCCTTTCCTGCAACCGGACATCTGACGACTTAAC 3109  
|||  
Db 3236 ATTGCTGTGAGGACATCTCTCTGAGCCTTTCCTGCAACCGGACATCTGACGACTTAAC 3385  
|||  
Qy 3110 TGTGCAAGATTAACCTTCAAGTCCCAAGAGATGATGAAGCTGTGTTGAGCTTGGCTGTC 3169  
|||  
Db 3386 TGTGCAAGATTAACCTTCAAGTCCCAAGAGATGATGAAGCTGTGTTGAGCTTGGCTGTC 3445  
|||  
Qy 3170 CCAAGTCTTAATTAACATTAATTTGGCTGTGGAAGATGAGCACTTCTGTGAATTAAGA 3229  
|||  
Db 3446 CCAAGTCTTAATTAACATTAATTTGGCTGTGGAAGATGAGCACTTCTGTGAATTAAGA 3505  
|||  
Qy 3220 AGCTGTGAGAGAGTGAAGTGAAGTCAAGCCCGAGTGTGTAATGAAGTGAAGTGAAGT 3289  
|||  
Db 3506 AGCTGTGAGAGAGTGAAGTGAAGTCAAGCCCGAGTGTGTAATGAAGTGAAGTGAAGT 3565  
|||  
Qy 3290 CTTTGTGATGAAGATGACCG 3308  
|||  
Db 3566 CTTTGTGATGAAGATGACCG 3584  
|||

RESULT 9  
AX427610  
LOCUS AX427610 3900 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 23 from Patent WO0232955.  
ACCESSION AX427610  
VERSION AX427610.1 GI:21537730  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Nelson, L.M. and Tong, Z.-B.  
TITLE Human gene critical to fertility  
JOURNAL Patent: WO 0232955-A 23 25-APR-2002;  
GOVERNMENT OF THE UNITED STATES (US)  
FEATURES  
location/Qualifiers  
source  
1..3900  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
1..3603  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD35973.1"  
/db\_xref="GI:21537731"



Qy	2090	GCAGCAGCATCTCTGACAAAGGGGGGCATGAAGA	CCCTGTGTGGCCAAAGCTGAGGCAATCCCA	214.9
Dp	2366	GCAGCAGCATCTCTGACAAAGGGGGCCATGAAGA	CCCTGTGTGGCCAAAGCTGAGGCAATCCCA	242.5
Qy	2150	CTTCGAAGATACAGACCCCTGATGTTTATGAAA	TGCACAGATTAACCCCTGTGTGACGACCC	220.9
Dp	2426	CCTCGAAGATACAGACCCCTGATGTTTATGAAA	TGCACAGATTAACCCCTGTGTGACGACCC	248.5
Qy	2210	TCTTGAGAAATGTCATGTGCCCAACCGTAACT	CTTAAGATCCCTCAACTTGGGAGGCAACCAAC	228.5
Dp	2486	TCTTGAGAAATGTCATGTGCCCAACCGTAACT	CTTAAGATCCCTCAACTTGGGAGGCAACCAAC	254.5
Qy	2270	TGAAGAAAGAGATGTATAAGATGGCGGTGAGAC	CTTTAAACCAACCAAAATGTTTGTGTGG	233.9
Dp	2546	TGAAGAAAGAGATGTATAAGATGGCGGTGAGAC	CTTTAAACCAACCAAAATGTTTGTGTGG	260.5
Qy	2330	AGTCTTTGAGGCTGGATGTCGTGTGATTTGACC	ATGACCATGCTGTATCTGAAGATCTTCCAAA	238.9
Dp	2606	AGTCTTTGAGGCTGGATGTCGTGTGATTTGACC	ATGACCATGCTGTATCTGAAGATCTTCCAAA	266.5
Qy	2390	TCTTTACGACCTCCCCCAGCCTGAATCTCTGAG	CTTGGCAGGAAACAAGTGCACAGACC	244.9
Dp	2666	TCTTTACGACCTCCCCCAGCCTGAATCTCTGAG	CTTGGCAGGAAACAAGTGCACAGACC	272.5
Qy	2450	AGGGAGTAATGCCTCTCAGTATGCTTGAAGAT	CTCCAGTGTGCCCTGTGAGAAAGCTGA	250.9
Dp	2726	AGGGAGTAATGCCTCTCAGTATGCTTGAAGAT	CTCCAGTGTGCCCTGTGAGAAAGCTGA	278.5
Qy	2510	TACTGAGAGACTGTGGCATCAACAGCACA	CGGGTTGCCAGAGATGTGGCCTCAAGCCTCTGCA	256.9
Dp	2786	TACTGAGAGACTGTGGCATCAACAGCACA	CGGGTTGCCAGAGATGTGGCCTCTGCA	284.5
Qy	2570	GCAACCGGAGCTTTGACACACCTGTGCTTAT	CCCAACAACAGCTGTGGGAAACGAAGTGTAA	262.9
Dp	2846	GCAACCGGAGCTTTGACACACCTGTGCTTAT	CCCAACAACAGCTGTGGGAAACGAAGTGTAA	290.5
Qy	2630	ATCTACTGTGTGCATTCATCAATGAGAGCTT	CCCCACTGATAGCTGCAGAGGCTGATGCTGAATC	268.9
Dp	2906	ATCTACTGTGTGCATTCATCAATGAGAGCTT	CCCCACTGATAGCTGCAGAGGCTGATGCTGAATC	296.5
Qy	2690	AGTGCACCTTGACAACCGGCTGGCTGTGTG	ATTTCTTGCACTTGCAGCTTAATGGGTAATCTCAT	274.9
Dp	2966	AGTGCACCTTGACAACCGGCTGGCTGTGTG	ATTTCTTGCACTTGCAGCTTAATGGGTAATCTCAT	302.5
Qy	2750	GGCTGACGCACTTGAGCCTTAGCATGAAC	CCTGTGTGAAGAACAATGGCGTGAAGCTTCTGT	280.9
Dp	3026	GGCTGACGCACTTGAGCCTTAGCATGAAC	CCTGTGTGAAGAACAATGGCGCTGAAGCTTCTGT	308.5
Qy	2810	GCGAGGTGATGAGAGAAACATCTTGTCATCT	CCACAGACCTGAGGTGGTAAAGTGCATC	286.9
Dp	3086	GCGAGGTGATGAGAGAAACATCTTGTCATCT	CCACAGACCTGAGGTGGTAAAGTGCATC	314.5
Qy	2870	TCACCGCGCGCTGTGTGAGAGTCTGTCT	CTGTGTGATCTCGAGAGACGACACCTGAAGA	292.9
Dp	3146	TCACCGCGCGCTGTGTGAGAGTCTGTCT	CTGTGTGATCTCGAGAGACGACACCTGAAGA	320.5
Qy	2930	GCTGTGATCTCAACGACCAATGCGCTGGGTGA	CGGTGGGCTTGTGTGCGCTGTGTGCAAGGAC	298.9
Dp	3206	GCTGTGATCTCAACGACCAATGCGCTGGGTGA	CGGTGGGCTTGTGTGCGCTGTGTGCAAGGAC	326.5
Qy	2990	TGAAGCAAAAAGAACAGTGTCTGACAGAC	CTCGGGTTGGAAGGCAATGTGTGATCTGAC	304.9
Dp	3266	TGAAGCAAAAAGAACAGTGTCTGACAGAC	CTCGGGTTGGAAGGCAATGTGTGATCTGAC	332.5
Qy	3050	ATTGCTGTGAGGCACTCTCTTTGGCCCTTTT	CTGTGCAACCGGACATCTGACCAAGTCTAAACC	310.9
Dp	3326	ATTGCTGTGAGGCACTCTCTTTGGCCCTTTT	CTGTGCAACCGGACATCTGACCAAGTCTAAACC	338.5
Qy	3110	TGTGTGACGAATTAATCTTCAGTCTCCAA	AGGAATGATGAAAGCTGTGTTCGGCCTTTTGTGCTGTCT	316.9
Dp	3386	TGTGTGACGAATTAATCTTCAGTCTCCAA	AGGAATGATGAAAGCTGTGTTCGGCCTTTTGTGCTGTCT	344.5

QY	3170	CCAGGTCTAACTTACAGATTAATTTGGGCTGTGGGAATATGGACAGTACCTCTGTGCATAATTAAGA	3229
Db	3446	CCAGGTCTAACTTACAGATTAATTTGGGCTGTGGGAATATGGACAGTACCTCTGTGCATAATTAAGA	3505
QY	3230	AGCTGCTGAGGAAGTGCAGCTACTCAAGCCCGAGTCTGTAATTAACGGTATTTGGCATT	3289
Db	3506	AGCTGCTGAGGAAGTGCAGCTACTCAAGCCCGAGTCTGTAATTAACGGTATTTGGCATT	3565
QY	3290	CTTTTGATGAAGATGACCG 3308	
Db	3566	CTTTTGATGAAGATGACCG 3584	
RESULT 10			
LOCUS	CQ731113	2753 bp	DNA linear PAT 03-FEB-2004
DEFINITION	Sequence 17047 from Patent WO02068579.		
ACCESSION	CQ731113		
VERSION	CQ731113.1	GI:42306827	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1		
JOURNAL	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.		
FEATURES	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof		
Source	Patent: WO 02068579-A 17047 06-SEP-2002; PE Corporation (NY) (US)		
ORIGIN	Location/Qualifiers		
	1..2753		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
Query Match	48.0%; Score 1936; DB 6; Length 2753;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 1936; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
QY	404	GACATGAGGATGACACATGGGATCTCAACAAAGTCAAGTATGACCAAAATTCGCTGAGAGG	463
Db	1	GACATGAGGATGACACATGGGATCTCAACAAAGTCAAGTATGACCAAAATTCGCTGAGAGG	60
QY	464	AGGATGTAACGTCGTAGTTTGGAAAAACATGCTGCTGACTGGCCGGAATGCAAACGTTGG	523
Db	61	AGGATGTAACGTCGTAGTTTGGAAAAACATGCTGCTGACTGGCCGGAATGCAAACGTTGG	120
QY	524	CTGGTCTTTGATTTCACAGCCGCTGGGCGCTTCGCGCTCGACCGTGTTCGACGGA	583
Db	121	CTGGTCTTTGATTTCACAGCCGCTGGGCGCTTCGCGCTCGACCGTGTTCGACGGA	180
QY	584	AGTCAGGAATTTGGGAATTCGCTCTTACCCAGAAAGATCGTGTGCTGGGCGCAAGTG	643
Db	181	AGTCAGGAATTTGGGAATTCGCTCTTACCCAGAAAGATCGTGTGCTGGGCGCAAGTG	240
QY	644	GACTCTACCAAGGAATGTTCTCTACGCTCTTCTCTCCCGTTAGAGAGATGCAGCGGA	703
Db	241	GACTCTACCAAGGAATGTTCTCTACGCTCTTCTCTCCCGTTAGAGAGATGCAGCGGA	300
QY	704	AGAAAGAGAGACGATGTCAACAGATTCATCTTCAGAGGAGTGGCCAGACTCCAGGCTCCG	763
Db	301	AGAAAGAGAGACGATGTCAACAGATTCATCTTCAGAGGAGTGGCCAGACTCCAGGCTCCG	360
QY	764	TGACGGAGATATGTGCCCGACAGAAAGGCTGTTGTTCAATCATTTGAAGGTTTGCATGACC	823
Db	361	TGACGGAGATATGTGCCCGACAGAAAGGCTGTTGTTCAATCATTTGAAGGTTTGCATGACC	420
QY	824	TGGGCTCTGTCCTCAACATGACCAAAAGCTCTGGCAAAAGCTGGGCTGAGAAAGAGGCTC	883
Db	421	TGGGCTCTGTCCTCAACATGACCAAAAGCTCTGGCAAAAGCTGGGCTGAGAAAGAGGCTC	480

QY 884 CGTTACCCCTATACGAGTCTGTGAGAGAGTCTCTGCTCCCTGAGTCTTCTTGATCG 943  
 DB 481 CGTTACCCCTATACGAGTCTGTGAGAGAGTCTCTGCTCCCTGAGTCTTCTTGATCG 540  
 QY 944 TCACGTCAGAGAGTGGGCAAGAGAGTCAAGTCAAGAGTCTGTCCTCCCTTACC 1003  
 DB 541 TCACGTCAGAGAGTGGGCAAGAGAGTCAAGTCAAGAGTCTGTCCTCCCTTACC 600  
 QY 1004 TGTATAGAGAGAAATCTCCGGGGAACAAAGAAATCACTTGTCTCTTGAAGCGGGATTTG 1063  
 DB 601 TGTATAGAGAGAAATCTCCGGGGAACAAAGAAATCACTTGTCTCTTGAAGCGGGATTTG 660  
 QY 1064 GTGAGCATCAGAAACAAAGAGTGGGTCATGATGAAACAAAGTGAAGTCTGACAC 1123  
 DB 661 GTGAGCATCAGAAACAAAGAGTGGGTCATGATGAAACAAAGTGAAGTCTGACAC 720  
 QY 1124 AGTGCAGAGTGGGCTCTCATCTGAGTGGGCTCTGAGCTGACAGAGCTGG 1183  
 DB 721 AGTGCAGAGTGGGCTCTCATCTGAGTGGGCTCTGAGCTGACAGAGCTGG 780  
 QY 1184 TGGGGAGAGCGTGGGCTCTTCAACAAAGCTGACAGAGCTGACAGCGCTTTTGTGT 1243  
 DB 781 TGGGGAGAGCGTGGGCTCTTCAACAAAGCTGACAGAGCTGACAGCGCTTTTGTGT 840  
 QY 1244 TTGATCAGTCAACCTCTGAGAGCGTGGTCCGCGCTGTCTCAATCTGAGAGAAAGTTG 1303  
 DB 841 TTGATCAGTCAACCTCTGAGAGCGTGGTCCGCGCTGTCTCAATCTGAGAGAAAGTTG 900  
 QY 1304 TCTCTAAGAGGCTTCTGCGGTATGAGTGGGAGAGTGAATGAGAAATGAGTGTG 1363  
 DB 901 TCTCTAAGAGGCTTCTGCGGTATGAGTGGGAGAGTGAATGAGAAATGAGTGTG 960  
 QY 1364 ATGAGAGAGCTTCAAGTTCAGAGAGTCTGAGAGTCTGAGTCTGTTTCA 1423  
 DB 961 ATGAGAGAGCTTCAAGTTCAGAGAGTCTGAGAGTCTGAGTCTGTTTCA 1020  
 QY 1424 TGAACATCTTCTCCAGACAGCTGAGAGTCAACCTTCTTCCACTGATG 1483  
 DB 1021 TGAACATCTTCTCCAGACAGCTGAGAGTCAACCTTCTTCCACTGATG 1080  
 QY 1484 TCCAGAGCTTCTGAGCGCTGTATGAGTGAAGGAGCTGAAATTCAGAGCTG 1543  
 DB 1081 TCCAGAGCTTCTGAGCGCTGTATGAGTGAAGGAGCTGAAATTCAGAGCTG 1140  
 QY 1544 TCTGCGCTCTGAGTGAAGAGAGAGTTCATGAGAGTTCATGAGAGCTTCC 1603  
 DB 1141 TCTGCGCTCTGAGTGAAGAGAGAGTTCATGAGAGTTCATGAGAGCTTCC 1200  
 QY 1604 ATATCACTCGCTTGAAGAGAGTTCATGAGAGTTCATGAGAGCTTGA 1663  
 DB 1201 ATATCACTCGCTTGAAGAGAGTTCATGAGAGTTCATGAGAGCTTGA 1260  
 QY 1664 GGGCACTGAGAGTCTGAGCGCTGCTCCCTGAGGAGTGAAGAGAGTCTG 1723  
 DB 1261 GGGCACTGAGAGTCTGAGCGCTGCTCCCTGAGGAGTGAAGAGAGTCTG 1320  
 QY 1724 ACTGAGTCTCTGAGTGAAGAGAGTTCATGAGAGTTCATGAGAGCTTGA 1783  
 DB 1321 ACTGAGTCTCTGAGTGAAGAGAGTTCATGAGAGTTCATGAGAGCTTGA 1380  
 QY 1784 CTTTCACTGCTTTCAGAGACTCAAGAGAGTTCATGAGAGCTTGA 1843  
 DB 1381 CTTTCACTGCTTTCAGAGACTCAAGAGAGTTCATGAGAGCTTGA 1440  
 QY 1844 TCCAGAGTGGGCTTCCAGTTCAGAGAGTTCATGAGAGTTCATGAGAGCTTGA 1903  
 DB 1441 TCCAGAGTGGGCTTCCAGTTCAGAGAGTTCATGAGAGTTCATGAGAGCTTGA 1500  
 QY 1904 TCCAGAGTGGGCTTCCAGTTCAGAGAGTTCATGAGAGTTCATGAGAGCTTGA 1963  
 DB 1501 TCCAGAGTGGGCTTCCAGTTCAGAGAGTTCATGAGAGTTCATGAGAGCTTGA 1560

QY 1964 ATGAGTCCGTCAGAGCATGCTGAGTCCCTGATGAGTGGGAGTAAAGACCTCATTTG 2023  
 DB 1561 ATGAGTCCGTCAGAGCATGCTGAGTCCCTGATGAGTGGGAGTAAAGACCTCATTTG 1620  
 QY 2024 AGAGCATGAGAGATTTCTGCTCATGCTTGGACCAACCACTTGGGAGCTGG 2083  
 DB 1621 AGAGCATGAGAGATTTCTGCTCATGCTTGGACCAACCACTTGGGAGCTGG 1680  
 QY 2084 ACTGGGAGAGCATCTTCAAGAGCGGCGCATGAGAGCTTGTGTCCAACTGAGGC 2143  
 DB 1681 ACTGGGAGAGCATCTTCAAGAGCGGCGCATGAGAGCTTGTGTCCAACTGAGGC 1740  
 QY 2144 ATCCACCTGAGATCAGACCCCTGATGTTAGAAATGACAGATTAACCTGAGTGGC 2203  
 DB 1741 ATCCACCTGAGATCAGACCCCTGATGTTAGAAATGACAGATTAACCTGAGTGGC 1800  
 QY 2204 AGCAGCTTGAAGATCTGTCAGAGCCGTAACCTTAAGATCCCTCAACTTGGAGGCA 2263  
 DB 1801 AGCAGCTTGAAGATCTGTCAGAGCCGTAACCTTAAGATCCCTCAACTTGGAGGCA 1860  
 QY 2264 CCCACCTGAGAGAGATTTAGATGAGTGGTGAAGCTTAAACCAACCAATGTT 2323  
 DB 1861 CCCACCTGAGAGAGATTTAGATGAGTGGTGAAGCTTAAACCAACCAATGTT 1920  
 QY 2324 TGTGAGTCTTGGAG 2339  
 DB 1921 TGTGAGTCTTGGAG 1936

RESULT 11  
 AC011470/c  
 LOCUS AC011470 157141 bp DNA linear PRI 15-JUL-2000  
 DEFINITION Homo sapiens chromosome 19 clone CTC-490M10, complete sequence.  
 ACCESSION AC011470  
 VERSION AC011470.5 GI:9211204  
 KEYWORDS HTG;  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
 TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 COMMENT  
 On Jul 15, 2000 this sequence version replaced gi:7690109.  
 Draft Sequence Produced by DOE Joint Genome Institute  
 www.jgi.doe.gov  
 Finishing Completed at Stanford Human Genome Center  
 www.sbgc.stanford.edu  
 Quality: Phrap Quality >=40 99.9% of Sequence;  
 Estimated Total Number of Errors is 0.1.  
 FEATURES  
 source 1. 157141  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /clone="CTC-490M10"  
 ORIGIN  
 Query Match 38.4%; Score 1549; DB 9; Length 157141;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1599; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY	402	AGGACATGGAGGTGACACATGGGATACAGAGCTCACGTGATACCAAAATTGCTGAGGA	461
Db	42001	AGGACATGGAGGTGACACATGGGCTACAAAGATCACGTGATACCAAAATTGCTGAGGA	41942
OY	462	GGAGAGTGTACGTGTAGTTTTGAAAACCTGCTGTGATCTGAGCCGGAAATGCAACGTT	521
Db	41941	GGAGAGTGTACGTGTAGTTTTGAAAACCTGCTGTGATCTGAGCCGGAAATGCAACGTT	41882
OY	522	GGCTGTGTCTTTTAAATTCACACCGGTGGGCTTCGGCTTCGACGGTGTCTGCACGG	581
Db	41881	GGCTGTGTCTTTTAAATTCACACCGGTGGGCTTCGGCTTCGACGGTGTCTGCACGG	41822
OY	582	AAAGTCAGGAATTTGGGAAATCGGCTCAGCCAAAGGATCGTCTGTGCGGCGCAAG	641
Db	41821	AAAGTCAGGAATTTGGGAAATCGGCTCAGCCAAAGGATCGTCTGTGCGGCGCAAG	41762
OY	642	TGACCTCACCAGGAAATGTTCTCTACGTCTTCTCTCCCGTTAGAGATGCACGG	701
Db	41761	TGACCTCACCAGGAAATGTTCTCTACGTCTTCTCTCCCGTTAGAGATGCACGG	41702
OY	702	GAAAGAGAGAGCAGTGTCAACAGATTCACTCCAGGAGTGGCAGACTCCCAAGCTCC	761
Db	41701	GAAAGAGAGAGCAGTGTCAACAGATTCACTCCAGGAGTGGCAGACTCCCAAGCTCC	41642
OY	762	GGTAGCAGAGATCAATGTCCTCCGACCAAGAGGCTGTGTTCAATTAAGCGTTTCATGA	821
Db	41641	GGTAGCAGAGATCAATGTCCTCCGACCAAGAGGCTGTGTTCAATTAAGCGTTTCATGA	41582
OY	822	CTGTGGGCTGTCTCCTCAACAATGACACAAAGCTCTCAAGATGGGCTGAGAGCAGCC	881
Db	41581	CTGTGGGCTGTCTCCTCAACAATGACACAAAGCTCTCAAGATGGGCTGAGAGCAGCC	41522
OY	882	TCCGTTCAACCTCATAGCGAGCTGTGCTGAGAGAGTCTGCTCTCCAGATCTTCTGAT	941
Db	41521	TCCGTTCAACCTCATAGCGAGCTGTGCTGAGAGAGTCTGCTCTCCAGATCTTCTGAT	41462
OY	942	CGTCAACGTCAGAGACGTGGCACAGAGAGCTCAAGTCAGAGGTGTGTCTTCCCGTTA	1001
Db	41461	CGTCAACGTCAGAGACGTGGCACAGAGAGCTCAAGTCAGAGGTGTGTCTTCCCGTTA	41402
OY	1002	CCTGTTAGTTAGAGAAATCTCCGGGAAACAAAGATTCATTGCTCTTTGAGGCGAGAT	1061
Db	41401	CCTGTTAGTTAGAGAAATCTCCGGGAAACAAAGATTCATTGCTCTTTGAGGCGAGAT	41342
OY	1062	TGCTGAGCATCAGAAACACAAAGGTTGCGTGCATATGAAACAACGTGAGCTGTCSA	1121
Db	41341	TGCTGAGCATCAGAAACACAAAGGTTGCGTGCATATGAAACAACGTGAGCTGTCSA	41282
OY	1122	CCAGTGCAGAGTGGCCGCGCGTGGGCTCTCTCATCTGTGTGGGCCCTGACGTGAGGAGCT	1181
Db	41281	CCAGTGCAGAGTGGCCGCGCGTGGGCTCTCTCATCTGTGTGGGCCCTGACGTGAGGAGCT	41222
OY	1182	GGTGGGGAGACGTGCGCCCTTCAACAAACGCTCAAGGCTTGCAACGCGCTTTTGT	1241
Db	41221	GGTGGGGAGAGCGTGCCTCCCTTCAACAAACGCTCAAGGCTTGCAACGCGCTTTTGT	41162
OY	1242	GTTTCATCAGTCAACCCCTGAGAGCGGTGTCCGCGCTGTCTCAACTGTGAGAGAAAGAT	1301
Db	41161	GTTTCATCAGTCAACCCCTGAGAGCGGTGTCCGCGCTGTCTCAACTGTGAGAGAAAGAT	41102
OY	1302	TGTCCTAAGCGCTTTCGCGTATGCTGTGAGAGGAGTGTGGAATAGGAAGTCAAGTGT	1361
Db	41101	TGTCCTAAGCGCTTTCGCGTATGCTGTGAGAGGAGTGTGGAATAGGAAGTCAAGTGT	41042
OY	1362	TGATGTGACGACCTCATGTTCAAGGATCGGGAGTCTGAGCTCCGTGCTGTGTTCA	1421
Db	41041	TGACGTGACGACCTCATGTTCAAGGATCGGGAGTCTGAGCTCCGTGCTGTGTTCA	40982
OY	1422	CATGAACATCTTCTCCCAAGACCACTGTGAGAGTACTCACTTCTTCAACCTCACTCAG	1481
Db	40981	CATGAACATCTTCTCCCAAGACCACTGTGAGAGTACTCACTTCTTCAACCTCACTCAG	40922

QY	1482	CTTCAGAGACTCTGTCGCGCCTTGTACTACCGTTAGTAGAGGGCCTGGAATCGAGCAC	1541
Db	40921	TTCTCAGAGACTTCTGTGCGCCTTGTACTACGTTAGTAGAGGGCCTGGAATCGAGCAC	40862
QY	1542	TTCTGCGCCTCTGTACGTTGAGAAACAAGAGTCCATGAGACTTTAAACAGCAGCTT	1601
Db	40861	TTCTGCGCCTCTGTACGTTGAGAAACAAGAGTCCATGAGACTTTAAACAGCAGCTT	40802
QY	1602	CCATATCCACTCCGCTTTGATGTAAGAGGTTCTTGTTTGACCTCGTAGAGGAAGAGTAG	1661
Db	40801	CCATATCCACTCCGCTTTGATGTAAGAGGTTCTTGTTTGACCTCGTAGAGGAAGAGTAG	40742
QY	1662	GAGGCCACTGAGAGTCCGTCGTGGGCTGTCCCTTCCCTGGGGGTGAACAGAACTTCT	1721
Db	40741	GAGGCCACTGAGAGTCCGTCGTGGGCTGTCCCTTCCCTGGGGGTGAACAGAACTTCT	40682
QY	1722	GCACTGGGTCCTCTGTGGGTGAGAGCCCTATATGCACACCCCGAGAGACACCCCTGGA	1781
Db	40681	GCACTGGGTCCTCTGTGGGTGAGAGCCCTATATGCACACCCCGAGAGACACCCCTGGA	40622
QY	1782	CGCCTTCCACTGTCTTTCGAGACTCAAGACAAGAGTTGTTCCGTTGGACTTAAACG	1841
Db	40621	CGCCTTCCACTGTCTTTCGAGACTCAAGACAAGAGTTGTTCCGTTGGACTTAAACG	40562
QY	1842	CTTCCAGAAGTGGGCTTCGATTAACAGAACTCGACTGATATGCATCTTCCCTCG	1901
Db	40561	CTTCCAGAAGTGGGCTTCGATTAACAGAACTCGACTGATATGCATCTTCCCTCG	40502
QY	1902	CTTCCAGACACTGTCCGATTTTGGCGAAATTCGGGTGATATGTCMAAGGATCTTCCCAAG	1961
Db	40501	CTTCCAGACACTGTCCGATTTTGGCGAAATTCGGGTGATATGTCMAAGGATCTTCCCAAG	40442
QY	1962	AGATGAGTCCGCTGAGGCATGTCTCTGTGGTCCCTCTATGG	2001
Db	40441	AGATGAGTCCGCTGAGGCATGTCTCTGTGGTCCCTCTATGG	40402

RESULT 12				
AC024580	LOCUS			
AC024580	DEFINITION	AC024580	193609 bp	DNA linear PRI 21-DEC-2001
AC024580	VERSION	AC024580.6	GI:17975240	Homo sapiens chromosome 19 clone CTD-262117, complete sequence.
	KEYWORDS	HTG;		
	SOURCE	Homo sapiens (human)		
	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	1 (bases 1 to 193609) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission Unpublished		
REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	2 (bases 1 to 193609) DOE Joint Genome Institute. Direct Submission		
REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	Submitted (29-FEB-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	3 (bases 1 to 193609) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission		
REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	Submitted (13-NOV-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	4 (bases 1 to 193609) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission		
REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	Submitted (21-DEC-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
COMMENT	COMMENT	On Dec 21, 2001 this sequence version replaced gi:16905144. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov		
		Finishing Completed at Stanford Human Genome Center www.sngc.stanford.edu		
		Quality: Phrap Quality >=40.99.4% of Sequence;		





Query Match 27.3%; Score 1102; DB 6; Length 1157;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 757 GCTCCGGAGGAGGATCATGTCCCGACCAAGAGGCTGTGTTCATCATTTGACGTTTC 816
    |||
DB 4 GCTCCGGAGGAGGATCATGTCCCGACCAAGAGGCTGTGTTCATCATTTGACGTTTC 63
OY 817 GATGACCTGGGCTGTCTCTCAACAATGACAAAGCTCTGCAAGAGCTGGGCTGAGAG 876
    |||
DB 64 GATGACCTGGGCTGTCTCTCAACAATGACAAAGCTCTGCAAGAGCTGGGCTGAGAG 123
OY 877 CAGCCTCCGTTACACCTCATACGACATCTGTGAGAGAGTCTGCTCCCTGAGTCTTTC 936
    |||
DB 124 CAGCCTCCGTTACACCTCATACGACATCTGTGAGAGAGTCTGCTCCCTGAGTCTTTC 183
OY 937 CTGATCTGACCTGTCAGAGAGCTGGGCAACAGAGAGTCTGAGAGTCTGCTTCC 996
    |||
DB 184 CTGATCTGACCTGTCAGAGAGCTGGGCAACAGAGAGTCTGAGAGTCTGCTTCC 243
OY 997 CGTTACCTGTTAGTTAGAGGAATCTCCGGGGAAACAAATTCATCTGCTCTTGAGGCG 1056
    |||
DB 244 CGTTACCTGTTAGTTAGAGGAATCTCCGGGGAAACAAATTCATCTGCTCTTGAGGCG 303
OY 1057 GGGATTGGTGACATCAGAAAGACAGAGGTTGCTGCGATCATGAAACCGTGAGCTG 1116
    |||
DB 304 GGGATTGGTGACATCAGAAAGACAGAGGTTGCTGCGATCATGAAACCGTGAGCTG 363
OY 1117 CTGACACAGTGCAGAGTCCCGCGTGGGCTCTCTCATCTGTGCGCTCTGAGTCTGAG 1176
    |||
DB 364 CTGACACAGTGCAGAGTCCCGCGTGGGCTCTCTCATCTGTGCGCTCTGAGTCTGAG 423
OY 1177 GACGCGGGGGGAGAGGGTGGCGCCCTTCAACCAAGCTCAGAGCTGACGCGCT 1236
    |||
DB 424 GACGCGGGGGGAGAGGGTGGCGCCCTTCAACCAAGCTCAGAGCTGACGCGCT 483
OY 1237 TTTGTGTTTCATCAGCTCACCCTCGAGGCGTGTCTCGGCGTGTCTCATCTGAGGAA 1296
    |||
DB 484 TTTGTGTTTCATCAGCTCACCCTCGAGGCGTGTCTCGGCGTGTCTCATCTGAGGAA 543
OY 1297 AGAGTTGTCGAGAGCGCTTCTGCGGTAATGCTGTGAGAGGAGTGTGAATGAGATCA 1356
    |||
DB 544 AGAGTTGTCGAGAGCGCTTCTGCGGTAATGCTGTGAGAGGAGTGTGAATGAGATCA 603
OY 1357 GTGTTGATGAGAGACCTCATGTTGAGAGACCTGGGGAGTCTGAGCTCCGCTCTG 1416
    |||
DB 604 GTGTTGATGAGAGACCTCATGTTGAGAGACCTGGGGAGTCTGAGCTCCGCTCTG 663
OY 1417 TTTCAATGAAATCTCTTCTCCCAAGACCACTGTGAGAGTACTAATCACTTCTTCCAC 1476
    |||
DB 664 TTTCAATGAAATCTCTTCTCCCAAGACCACTGTGAGAGTACTAATCACTTCTTCCAC 723
OY 1477 CTCACTCTCAGAGATCTCTGTCGCGCTTGTGTAAGTGTGAGAGGCGCTGGAATTCAG 1536
    |||
DB 724 CTCACTCTCAGAGATCTCTGTCGCGCTTGTGTAAGTGTGAGAGGCGCTGGAATTCAG 783
OY 1537 CCAGCTCTCTGCGCTCTGAGAGTGAAGAGCAAGAGAGTCCATGAGCTTAAACAGGCA 1596
    |||
DB 784 CCAGCTCTCTGCGCTCTGAGAGTGAAGAGCAAGAGAGTCCATGAGCTTAAACAGGCA 843
OY 1597 GAGCTTCATATCACTCGCTTGTGATGAAGGCTTCTTGTGTGAGCTGTCGTCGAGAAC 1656
    |||
DB 844 GAGCTTCATATCACTCGCTTGTGATGAAGGCTTCTTGTGTGAGCTGTCGTCGAGAAC 903
OY 1657 GTAAGAGAGCCACTGAGAGTCTGTCGAGCTGTCGCTTCCCTGAGGAGTGAAGCAAG 1716
    |||
DB 904 GTAAGAGAGCCACTGAGAGTCTGTCGAGCTGTCGCTTCCCTGAGGAGTGAAGCAAG 963
OY 1717 CTTCGCACTGGGCTCTCTGTGGGTGAGAGCTTAATGACACCAACCCAGAGACACC 1776
    |||
DB 964 CTTCGCACTGGGCTCTCTGTGGGTGAGAGCTTAATGACACCAACCCAGAGACACC 1023
OY 1777 CTGAGAGCTTCACATGCTTTTTCGAGAGCTCAAGACAAAGATTGTCTGCTGGCATTA 1836
  
```

```

DB 1024 CTGAGAGCTTCACATGCTTTTCGAGACTCAAGACAAAGATTGTCTGCTGGCATTA 1083
OY 1837 AACAGCTTCGAAGAGTGTGCTTCGATTAAACAGAACTGAGCTGATGACATCTTCC 1896
DB 1084 AACAGCTTCGAAGAGTGTGCTTCGATTAAACAGAACTGAGCTGATGACATCTTCC 1143
OY 1897 TTCTGCTTCAGC 1909
DB 1144 TTCTGCTTCAGC 1156

```

RESULT 14  
 AC012107/c  
 LOCUS  
 DEFINITION Homo sapiens clone Rpl1-45K21, WORKING DRAFT SEQUENCE, 25 unordered  
 pieces.  
 ACCESSION AC012107.2 GI:7329252  
 VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 167509)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 Homo sapiens, clone Rpl1-45K21  
 Unpublished  
 2 (bases 1 to 167509)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,  
 Baldwin, J., Barna, N., Beckert, R., Boguslavsky, L., Boukhalter, B.,  
 Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,  
 Cooke, P., Dekrellano, K., Dewar, K., Domingo, M., Donelan, L., Doyle, M.,  
 Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,  
 Galagan, D., Gardyna, S., Grant, G., Hagos, B., Harford, A., Horton, L.,  
 Howland, D., C., Johnson, R., Jones, C., Kahn, L., Karate, A., Klein, J.,  
 Lehoczek, J., Lien, C., Locke, R., Macdonald, P., Marquis, A.,  
 McMan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,  
 Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
 Stange-Thomann, N., Stojanovic, N., Sudrmanian, A., Talamas, D.,  
 Teste, S., Tittell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
 Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (20-OCT-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 26, 2000 this sequence version replaced gi:6088020.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: M1BR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L1180  
 Center clone name: 45\_K21  
 ----- Summary Statistics  
 Sequencing vector: M13; M77815; 100% of reads  
 Chemistry: Dye-terminator-amersham; 4% of reads  
 Chemistry: Dye-terminator Big Dye; 96% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 145437 bases at least Q40  
 Consensus quality: 155496 bases at least Q30  
 Consensus quality: 159832 bases at least Q20  
 Insert size: 176000; agarose-fp  
 Insert size: 165109; sum-of-contigs  
 Quality coverage: 3.8 in Q20 bases; agarose-fp  
 Quality coverage: 4.1 in Q20 base.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 25 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is



- \* arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
- \* This record will be updated with the finished sequence as soon as it is available and the accession number will
- \* be preserved.

1	1061:	contig of 1061 bp	in length
*	1161:	gap of 100 bp	
*	1162	2381:	contig of 1220 bp in length
*	2382	2481:	gap of 100 bp
*	2482	3771:	contig of 1290 bp in length
*	3772	3871:	gap of 100 bp
*	3872	4949:	contig of 1078 bp in length
*	4950	5049:	gap of 100 bp
*	5050	6119:	contig of 1070 bp in length
*	6120	6219:	gap of 100 bp
*	6220	8023:	contig of 1804 bp in length
*	8024	8123:	gap of 100 bp
*	8124	10555:	contig of 2432 bp in length
*	10556	10655:	gap of 100 bp
*	10556	12768:	contig of 2113 bp in length
*	12769	12868:	gap of 100 bp
*	12869	15722:	contig of 2854 bp in length
*	15723	15822:	gap of 100 bp
*	15823	17959:	contig of 2137 bp in length
*	17960	18059:	gap of 100 bp
*	18060	21771:	contig of 3712 bp in length
*	21772	21871:	gap of 100 bp
*	21872	24539:	contig of 2668 bp in length
*	24540	24639:	gap of 100 bp
*	24640	28144:	contig of 3505 bp in length
*	28145	28244:	gap of 100 bp
*	28245	31503:	contig of 3259 bp in length
*	31504	31603:	gap of 100 bp
*	31604	35856:	contig of 4255 bp in length
*	35859	35958:	gap of 100 bp
*	35959	39610:	contig of 3552 bp in length
*	39611	39710:	gap of 100 bp
*	39711	42759:	contig of 3049 bp in length
*	42760	42859:	gap of 100 bp
*	42860	48014:	contig of 5155 bp in length
*	48015	48114:	gap of 100 bp
*	48115	52265:	contig of 4149 bp in length
*	52266	52363:	gap of 100 bp
*	52364	56376:	contig of 4011 bp in length
*	56375	56474:	gap of 100 bp
*	56475	63102:	contig of 6628 bp in length
*	63103	63202:	gap of 100 bp
*	63203	72687:	contig of 9485 bp in length
*	72688	72787:	gap of 100 bp
*	72789	97899:	contig of 25012 bp in length
*	97900	97989:	gap of 100 bp
*	97900	134325:	contig of 36426 bp in length
*	134326	134425:	gap of 100 bp
*	134426	167509:	contig of 33084 bp in length

```
misc_feature      8124.  .10555
                   /note="assembly_fragment "
misc_feature      10566.  .12768
                   /note="assembly_fragment "
misc_feature      12869.  .15722
                   /note="assembly_fragment "
misc_feature      15823.  .17959
                   /note="assembly_fragment "
misc_feature      18060.  .21771
                   /note="assembly_fragment "
misc_feature      21872.  .24539
                   /note="assembly_fragment "
misc_feature      24640.  .28144
                   /note="assembly_fragment "
misc_feature      28245.  .31503
                   /note="assembly_fragment "
misc_feature      31604.  .35858
                   /note="assembly_fragment "
misc_feature      35599.  .39610
                   /note="assembly_fragment "
misc_feature      42860.  .48014
                   /note="assembly_fragment "
misc_feature      48115.  .52263
                   /note="assembly_fragment "
misc_feature      52364.  .55374
                   /note="assembly_fragment "
misc_feature      56475.  .63102
                   /note="assembly_fragment "
misc_feature      63203.  .72687
                   /note="assembly_fragment "
```

ORIGIN

Query Match	22.3%	Score 900;	DB 2;	Length 167509;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 900;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1102	AAACAACGTAGAGTGCCTCGACACAGTGCACAGTGTCCGCGGTGGGGCTCTCAATCTGGCG	1161
Db	12664	AAACAACGTAGAGTGCCTCGACACAGTGCACAGTGTCCGCGGTGGGGCTCTCAATCTGGCG	12605
QY	1162	GCCCTGCAGCTGCAGGACGTGTGTGGGGAGAGCGTGCCTCTCAACCAACGCTCAC	1221
Db	12604	GCCCTGCAGCTGCAGGACGTGTGTGGGGAGAGCGTGCCTCTCAACCAACGCTCAC	12545
QY	1222	GGCCTGCACGCGCTTTTGTGTTTCATCAGCTCACCCCTGAAGCGCTGTCCGGCGCTGT	1281
Db	12544	GGCCTGCACGCGCTTTTGTGTTTCATCAGCTCACCCCTGAAGCGCTGTCCGGCGCTGT	12485
QY	1282	CTCAATCTGGAGGAAAGATTGTCTCGAAGCGCTTCTGCCGTATGGCTGTGGAGGAGTGC	1341
Db	12484	CTCAATCTGGAGGAAAGATTGTCTCGAAGCGCTTCTGCCGTATGGCTGTGGAGGAGTGC	12425
QY	1342	TGGAATGGAAGTCAAGTGTGTGATGGTGAAGACTCATGTGTTCAAGGACTCGGGGAGTGT	1401
Db	12424	TGGAATGGAAGTCAAGTGTGTGATGGTGAAGACTCATGTGTTCAAGGACTCGGGGAGTGT	12365
QY	1402	GAGCTCCGTGCTCTGTTTCAATGAACATCTCTTCCAGACAGCCACTGTGAGGAGTAC	1461
Db	12364	GAGCTCCGTGCTCTGTTTCAATGAACATCTCTTCCAGACAGCCACTGTGAGGAGTAC	12305
QY	1462	TACACTTCTTCCACCTCAAGTCTCCAGGACTTCTGTGCCGCTTGTAAGCTTTAGAG	1521

Db 12304 TACACCTTCCACCTCAGTCTCCAGGACCTCTGCGCCCTTGTATCACTGTTAGAG 12245  
Qy 1522 GGCCTGGAATCGAGCCAGCTCTCTGCCCCTCTGTACGTTGAGAAACAAGAGCTCATG 1581  
Db 12244 GGCCTGGAATCGAGCCAGCTCTCTGCCCCTCTGTACGTTGAGAAACAAGAGCTCATG 12185  
Qy 1582 GAGCTTAAACAGAGCGCTTCATTCATTCAGTCTGCTTGTGATGAAAGCTTTCTGTTTGGC 1641  
Db 12184 GAGCTTAAACAGAGCGCTTCATTCATTCAGTCTGCTTGTGATGAAAGCTTTCTGTTTGGC 12125  
Qy 1642 CTCGTAGCAGAGAGCTTGAAGAGGCGCTGAGAGTCTGCTGAGGCTGCTCCGTTCCCTG 1701  
Db 12124 CTCGTAGCAGAGAGCTTGAAGAGGCGCTGAGAGTCTGCTGAGGCTGCTCCGTTCCCTG 12065  
Qy 1702 GGGGTGAAGCAGAGAGCTTCTGCACTGAGGTCTCTGTTGGGTGAGAGCTTAAATGCCACC 1761  
Db 12064 GGGGTGAAGCAGAGAGCTTCTGCACTGAGGTCTCTGTTGGGTGAGAGCTTAAATGCCACC 12005  
Qy 1762 ACCCGAGAGACACCTGAGAGCGCTTCCACTGCTTTTTCAGAGCTCAAGACAAAGAGTTT 1821  
Db 12004 ACCCGAGAGACACCTGAGAGCGCTTCCACTGCTTTTTCAGAGCTCAAGACAAAGAGTTT 11945  
Qy 1822 GTTGTGCTGCAATTAACAGCTTCCAGAGGTGCTTCCGATTACCAAGACCTGGAC 1881  
Db 11944 GTTGTGCTGCAATTAACAGCTTCCAGAGGTGCTTCCGATTACCAAGACCTGGAC 11885  
Qy 1882 TTGATAGATCTTCTTCTGCTCTCCAGACATGCTCCGATTATTCGGAATAATCGGTGAT 1941  
Db 11884 TTGATAGATCTTCTTCTGCTCTCCAGACATGCTCCGATTATTCGGAATAATCGGTGAT 11825  
Qy 1942 GTCAAGAGGATCTTCCAGAGATGATGCTCGCTGAGGATGCTGCTGATGCTCTATAG 2001  
Db 11824 GTCAAGAGGATCTTCCAGAGATGATGCTCGCTGAGGATGCTGCTGATGCTCTATAG 11765

RESULT 15  
AX427590 1075 bp DNA linear PAT 20-JUN-2002

LOCUS DEFINITION Sequence 3 from Patent WO0232955.  
AX427590  
ACCESSION AX427590.1 GI:21537712

VERSION AX427590.1 GI:21537712  
KEYWORDS  
SOURCE

ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 Nelson, L.M. and Tong, Z.-B.  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS Human gene critical to fertility  
TITLE Patent: WO 0232955-A 3 25-APR-2002;  
JOURNAL GOVERNMENT OF THE UNITED STATES (US)

FEATURES  
Location/Qualifiers  
1..1075  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 16.3%; Score 657; DB 6; Length 1075;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 2550 TCTGCTCAGCCCTCGTCAGCAACCGAGCTTGACACACTGTGCTTATCCAAACAAG 2609  
Qy 1 TCTGCTCAGCCCTCGTCAGCAACCGAGCTTGACACACTGTGCTTATCCAAACAAG 60

Db 2610 CCTGGGAAACGAAAGGTAAATCTACTGTGATCCATGAGGCTTCCCACTGATGCT 2669  
Qy 61 CCTGGGAAACGAAAGGTAAATCTACTGTGATCCATGAGGCTTCCCACTGATGCT 120

Db 2670 GCAGAGCTGATGCTGAATCAGTGCACCTGAGACAGGCTGCTGCTTCTTGAAGT 2729  
Qy 1 GCAGAGCTGATGCTGAATCAGTGCACCTGAGACAGGCTGCTGCTTCTTGAAGT 120

Db 121 GCAGAGCTGATGCTGAATCAGTGCACCTGAGACAGGCTGCTGCTTCTTGAAGT 180  
Qy 2730 TGGCTTATGGGTAACTCATGGCTGAGCACTGAGCTTATGATGAACCTGTGAGAA 2789  
Db 181 TGGCTTATGGGTAACTCATGGCTGAGCACTGAGCTTATGATGAACCTGTGAGAA 240  
Qy 2790 CAATGGGTAAGAGCTTGTGCGAGGTGATGAGAAACAATCTGTGATCTCCAGAGCT 2849  
Db 241 CAATGGGTAAGAGCTTGTGCGAGGTGATGAGAAACAATCTGTGATCTCCAGAGCT 300  
Qy 2850 GGAAGTGTAAAGTGTATCTCACCGCGCTGCTGTGAGAGTGTGCTGTGATCTC 2909  
Db 301 GGAAGTGTAAAGTGTATCTCACCGCGCTGCTGTGAGAGTGTGCTGTGATCTC 360  
Qy 2910 GAGGAGCAGACCTGAAAGAGCTGATCTCACGGAACAATGCTGAGGTGAGGAGGT 2969  
Db 361 GAGGAGCAGACCTGAAAGAGCTGATCTCACGGAACAATGCTGAGGTGAGGAGGT 420  
Qy 2970 TGTGCGCTGTGCGAGGAGCTGAAAGCAAAAGAACAGTGTCTGACGAGACTCGGATTGAA 3029  
Db 421 TGTGCGCTGTGCGAGGAGCTGAAAGCAAAAGAACAGTGTCTGACGAGACTCGGATTGAA 480  
Qy 3030 GGCATGTGAGACTGATCTGATGCTGTGAGGCACTCTCTGCGCTTCTCTCAACCG 3089  
Db 481 GGCATGTGAGACTGATCTGATGCTGTGAGGCACTCTCTGCGCTTCTCTCAACCG 540  
Qy 3090 GCATGTGAGACTGATCTGATGCTGTGAGGCACTCTCTGCGCTTCTCTCAACCG 3149  
Db 541 GCATGTGAGACTGATCTGATGCTGTGAGGCACTCTCTGCGCTTCTCTCAACCG 600  
Qy 3150 GTGTTGCGCTTGTGCTGCTCCAGCTGCTAACTTACAGATTAATGGGCTGTGAAATGCA 3209  
Db 601 GTGTTGCGCTTGTGCTGCTCCAGCTGCTAACTTACAGATTAATGGGCTGTGAAATGCA 660  
Qy 3210 GTACCTGTGCAATTAAGAAAGCTGTGAGAGAAAGTCAAGCTCAGTACTCAAGCCCGAGTCT 3269  
Db 661 GTACCTGTGCAATTAAGAAAGCTGTGAGAGAAAGTCAAGCTCAGTACTCAAGCCCGAGTCT 720  
Qy 3270 AATTGAGGTAAGTGGGATCTTGTGATGAAGATGACCG 3308  
Db 721 AATTGAGGTAAGTGGGATCTTGTGATGAAGATGACCG 759

Search completed: July 19, 2005, 07:45:13  
Job time : 17013 secs



PT diagnosing stress-related, apoptotic and inflammatory responses, or for  
PT treating inflammatory and immune system disorders, cancers, or  
PT neurological diseases.

XX Claim 4, Fig 3, 167bp, English.

PS The invention comprises the amino acid and coding sequences of human  
CC PYRIN protein. The PYRIN protein and DNA sequences of the invention are  
CC useful for modulating and diagnosing stress-related, apoptotic and  
CC inflammatory responses. The PYRIN protein and DNA sequences are useful  
CC for treating: inflammatory disorders and immune system disorders (e.g.  
CC Crohn's disease, reactive arthritis, multiple sclerosis, contact  
CC dermatitis, psoriasis, graft rejection, allergies, viral infections and  
CC bacterial infections); cancer (e.g. leukemia); autoimmune disorders  
CC (e.g. systemic lupus erythematosus and arthritis); and neurological  
CC diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN  
CC protein and DNA sequences may also be used in screening assays, detection  
CC assays (e.g. chromosomal mapping, tissue typing or forensic biology),  
CC predictive medicine (e.g. diagnostic assays, clinical trials and  
CC pharmacogenomics) and transcription profiling. The present DNA sequence  
CC encodes the human PYRIN-5 protein

XX Sequence 4035 BP; 935 A; 1022 C; 1175 G; 903 T; 0 U; 0 Other;

Query Match 100.0%; Score 4035; DB 6; Length 4035;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 ATGAGAGAGACAAATTCGCTCACTTTTCAGCTACGGGCTGCAATGCTCTATAG 60
DB 1 ATGAGAGAGACAAATTCGCTCACTTTTCAGCTACGGGCTGCAATGCTCTATAG 60
OY 61 CTAGACAAGAGAAATTCAGACATTCAGAAATTAAGAAAGAAATCTTCAAGATCG 120
DB 61 CTAGACAAGAGAAATTCAGACATTCAGAAATTAAGAAAGAAATCTTCAAGATCG 120
OY 121 ACCACATGCTCTATTCACAGATTTGAAATCGAAGATGCCAAGTGAATGTCGCACTC 180
DB 121 ACCACATGCTCTATTCACAGATTTGAAATCGAAGATGCCAAGTGAATGTCGCACTC 180
OY 121 ACCACATGCTCTATTCACAGATTTGAAATCGAAGATGCCAAGTGAATGTCGCACTC 180
DB 121 ACCACATGCTCTATTCACAGATTTGAAATCGAAGATGCCAAGTGAATGTCGCACTC 180
OY 181 CTCTTGATGATATTAATGAGACATGCTGCTGGGCTGCTGCTCATTAAGCATCTTTGAA 240
DB 181 CTCTTGATGATATTAATGAGACATGCTGCTGGGCTGCTGCTCATTAAGCATCTTTGAA 240
OY 181 CTCTTGATGATATTAATGAGACATGCTGCTGGGCTGCTGCTCATTAAGCATCTTTGAA 240
DB 181 CTCTTGATGATATTAATGAGACATGCTGCTGGGCTGCTGCTCATTAAGCATCTTTGAA 240
OY 241 AACCAAGACCTGCGAACCCTCTCGGAGAAAGCAAGGATGACATGAAAAAATTTCAAA 300
DB 241 AACCAAGACCTGCGAACCCTCTCGGAGAAAGCAAGGATGACATGAAAAAATTTCAAA 300
OY 301 GCTATGAAACAAGAGTGCACAGACAGACAGACAGAAACAAGAAATTTCAACAGCT 360
DB 301 GCTATGAAACAAGAGTGCACAGACAGACAGACAGAAACAAGAAATTTCAACAGCT 360
OY 361 ATGGAACAAGAGTGCACAGACAGACAGACAGAAACAAGAAATTTCAACAGCT 420
DB 361 ATGGAACAAGAGTGCACAGACAGACAGACAGAAACAAGAAATTTCAACAGCT 420
OY 421 TGGGACTCAAGAGTCAAGTGAATGACAAATTCGCTGAGAGAGAGATGATCGTGTAGT 480
DB 421 TGGGACTCAAGAGTCAAGTGAATGACAAATTCGCTGAGAGAGAGATGATCGTGTAGT 480
OY 481 TTTGAAAACACTGCTGCTGACCTGCGACGGTGTCTTGCAACGAAAGTCAAGAAATTTGGAAA 540
DB 481 TTTGAAAACACTGCTGCTGACCTGCGACGGTGTCTTGCAACGAAAGTCAAGAAATTTGGAAA 540
OY 541 GACCGGTGGGCTTCCGGCTTCGCAACGGTGTCTTGCAACGAAAGTCAAGAAATTTGGAAA 600
DB 541 GACCGGTGGGCTTCCGGCTTCGCAACGGTGTCTTGCAACGAAAGTCAAGAAATTTGGAAA 600
OY 601 TCGGCTTACGCAAGAGATGCTGCTGAGGAGGAGCAAGTGAATCTTACAGGGAATG 660
DB 601 TCGGCTTACGCAAGAGATGCTGCTGAGGAGGAGCAAGTGAATCTTACAGGGAATG 660
OY 661 TTCTCTACGCTTCTCTCTCCCGTTAGAGAGATGACAGCGGAAGAAAGAGCAGTGTG 720
```

```
DB 661 TTCTCTACGCTTCTCTCTCCCGTTAGAGAGATGACAGCGGAAGAAAGAGCAGTGTG 720
OY 721 ACAGAGTTCATCTCAGAGAGTGGCCAGACTCCCAAGGCTCCGGTACGAGAGATCATGTC 780
DB 721 ACAGAGTTCATCTCAGAGAGTGGCCAGACTCCCAAGGCTCCGGTACGAGAGATCATGTC 780
OY 781 CGACAGAGAAAGGCTGTTGTTTCATCATTAAGACGTTTTCATGATGACGAGCTGTCTCTCAAC 840
DB 781 CGACAGAGAAAGGCTGTTGTTTCATCATTAAGACGTTTTCATGATGACGAGCTGTCTCTCAAC 840
OY 841 AATGACCAAAAGCTCTGCAAAAGACTGAGAGAGAGAGAGAGCTCCGTTACCTCATACGC 900
DB 841 AATGACCAAAAGCTCTGCAAAAGACTGAGAGAGAGAGAGAGCTCCGTTACCTCATACGC 900
OY 901 AGTGTGCTGAGAGAGAGTCTGCTCTGAGAGCTCTTCTGATTCGACACCTGACAGAGCTG 960
DB 901 AGTGTGCTGAGAGAGAGTCTGCTCTGAGAGCTCTTCTGATTCGACACCTGACAGAGCTG 960
OY 961 GGCAAGAGAGAGCTCAAGTCAAGAGTGGTCTCCCGTTACCTGTTAGTTAGAGAAATC 1020
DB 961 GGCAAGAGAGAGCTCAAGTCAAGAGTGGTCTCCCGTTACCTGTTAGTTAGAGAAATC 1020
OY 1021 TCCGGGAAACAAGATTCACATTTGAGCGCGGATTTGAGATGAGACAGAGACA 1080
DB 1021 TCCGGGAAACAAGATTCACATTTGAGCGCGGATTTGAGATGAGACAGAGACA 1080
OY 1081 CAAGGTTGCGTGGGATCATTAACAACCGTGAAGTGTCTCCACAGTCCAGGTCCCGGCC 1140
DB 1081 CAAGGTTGCGTGGGATCATTAACAACCGTGAAGTGTCTCCACAGTCCAGGTCCCGGCC 1140
OY 1141 GTGGGCTCTCTCATCTGGGAGTGGCCCTGACGCTGACAGACGAGTGGGGAGAGAGTGGCC 1200
DB 1141 GTGGGCTCTCTCATCTGGGAGTGGCCCTGACGCTGACAGACGAGTGGGGAGAGAGTGGCC 1200
OY 1201 CCTTTCAACAAGAGCTTCACAGGCTGACGCGCTGCTTGTGTTTCATCAAGCTCAACCTT 1260
DB 1201 CCTTTCAACAAGAGCTTCACAGGCTGACGCGCTGCTTGTGTTTCATCAAGCTCAACCTT 1260
OY 1261 CGAGGCGTGTGCGGCGCTGCTCAATCTGAGAGAAAGATGTCGTAAGCGCTTCTGCG 1320
DB 1261 CGAGGCGTGTGCGGCGCTGCTCAATCTGAGAGAAAGATGTCGTAAGCGCTTCTGCG 1320
OY 1321 CGTATGCTGTGAGAGAGTGTGAATAGAAAGTCAAGTGTGATGATGATGATGATGATGATG 1380
DB 1321 CGTATGCTGTGAGAGAGTGTGAATAGAAAGTCAAGTGTGATGATGATGATGATGATGATG 1380
OY 1381 GTTCAAGAGCTCGGAGAGTCTGAGCTCGTCTGTTTCATGATGAACATCTTCTCCCA 1440
DB 1381 GTTCAAGAGCTCGGAGAGTCTGAGCTCGTCTGTTTCATGATGAACATCTTCTCCCA 1440
OY 1441 GACAGCCTCTGAGAGATCTACACTTCTTCCACCTCAAGTCTCCAGAGATTTCTGAGCC 1500
DB 1441 GACAGCCTCTGAGAGATCTACACTTCTTCCACCTCAAGTCTCCAGAGATTTCTGAGCC 1500
OY 1501 GCTTGTATCTACGTTTATGAGAGGCTGAAATCGAGCAGGCTCTGCTCCCTGTAGCTT 1560
DB 1501 GCTTGTATCTACGTTTATGAGAGGCTGAAATCGAGCAGGCTCTGCTCCCTGTAGCTT 1560
OY 1561 GAGAAACAAGAGTCCATGAGAGCTTAAACAGGAGGCTTCCATATCACTGCTTTGG 1620
DB 1561 GAGAAACAAGAGTCCATGAGAGCTTAAACAGGAGGCTTCCATATCACTGCTTTGG 1620
OY 1621 ATGAAGGCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
DB 1621 ATGAAGGCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
OY 1681 CTGGGCTGTCCCGTTCCCTGAGGAGTGAAGAGAGAGCTTCTGCACTGAGGCTCTCTGTTG 1740
DB 1681 CTGGGCTGTCCCGTTCCCTGAGGAGTGAAGAGAGAGCTTCTGCACTGAGGCTCTCTGTTG 1740
OY 1741 GGTGAGAGAGCTTAATGCAACCAACCAAGAGAGACCTTGAAGGCTTCACTGCTTTTC 1800
```

Db 1741 GGTGACAGACCTAATGCCACACCCAGAGAGACCCCTGAGCGCTTCCACTGTCTTTTC 1800  
Qy 1801 GAGACTCAAGAACAAAGAGTTTGTGGCTTGGCATTTAAACAGTTTCCAAAGAGTGTGGCTT 1860  
Db 1801 GAGACTCAAGAACAAAGAGTTTGTGGCTTGGCATTTAAACAGTTTCCAAAGAGTGTGGCTT 1860  
Qy 1861 CCGATTAAACAGAAACCTGGACTTGTAGCATCTTCTCTGCTGCTGACAGACTGTCCGAT 1920  
Db 1861 CCGATTAAACAGAAACCTGGACTTGTAGCATCTTCTCTGCTGCTGACAGACTGTCCGAT 1920  
Qy 1921 TTGCGGAAAATTCCGGGTGATGTCAAGGATCTTCCAAAGATGAGTCCGCTGAGCA 1980  
Db 1921 TTGCGGAAAATTCCGGGTGATGTCAAGGATCTTCCAAAGATGAGTCCGCTGAGCA 1980  
Qy 1981 TGTCTGTGTGTCCTCTATGTATGTGCGGATTAAGACCTCTCATTTGAGAGAGAGTGGGAAAT 2040  
Db 1981 TGTCTGTGTGTCCTCTATGTATGTGCGGATTAAGACCTCTCATTTGAGAGAGAGTGGGAAAT 2040  
Qy 2041 TTCTGCTCATGCTTGGGACCCACCCACACCTGGGGAGCTGGACCTGGGGAGAGCATC 2100  
Db 2041 TTCTGCTCATGCTTGGGACCCACCCACACCTGGGGAGCTGGACCTGGGGAGAGCATC 2100  
Qy 2101 CTGACAGACGGGGCCATGAGAACCTGTGTGCAAGCTGAGGACATCCACCTGCAAGATA 2160  
Db 2101 CTGACAGACGGGGCCATGAGAACCTGTGTGCAAGCTGAGGACATCCACCTGCAAGATA 2160  
Qy 2161 CAGACCTGTATGTTTGAAGATGACAGATTACCTCTGTGTGACAGACTCTGTGAATA 2220  
Db 2161 CAGACCTGTATGTTTGAAGATGACAGATTACCTCTGTGTGACAGACTCTGTGAATA 2220  
Qy 2221 GTCATGGGCAACCGTAACTTAAGATCCCTCAACTTGGGAGGACCCACCTGAAAGAGAG 2280  
Db 2221 GTCATGGGCAACCGTAACTTAAGATCCCTCAACTTGGGAGGACCCACCTGAAAGAGAG 2280  
Qy 2281 GATGTAAAGATGCGCTGTGAAGCTTTAAACACCCAAATGTTTGTGAGTCTTTGAGG 2340  
Db 2281 GATGTAAAGATGCGCTGTGAAGCTTTAAACACCCAAATGTTTGTGAGTCTTTGAGG 2340  
Qy 2341 CTGATTCCTGTGATGATGACCCATCTGTGTAAGATCTCCAAATCTTTCAGACC 2400  
Db 2341 CTGATTCCTGTGATGATGACCCATCTGTGTAAGATCTCCAAATCTTTCAGACC 2400  
Qy 2401 TCCCCAGCCTGAAATCTCTGAGCCTGGCAGAGAAACAAGTGAACAGACAGAGGAATATG 2460  
Db 2401 TCCCCAGCCTGAAATCTCTGAGCCTGGCAGAGAAACAAGTGAACAGACAGAGGAATATG 2460  
Qy 2461 CCTCTCAGTATGCTTGAAGATCTCCAGTGCCTGACAGAGCTGATCTGAGAGAG 2520  
Db 2461 CCTCTCAGTATGCTTGAAGATCTCCAGTGCCTGACAGAGCTGATCTGAGAGAG 2520  
Qy 2521 TGTGGCATCAAGCAGCGGTTGCCAGAGTCTGAGCCTGACGCCCTGTGACAGACCGGAGC 2580  
Db 2521 TGTGGCATCAAGCAGCGGTTGCCAGAGTCTGAGCCTGACGCCCTGTGACAGACCGGAGC 2580  
Qy 2581 TTGAACACCTGTGCTATCCAAACAACCGCTGGGGAAGAGGTGAATCTTACTGTGT 2640  
Db 2581 TTGAACACCTGTGCTATCCAAACAACCGCTGGGGAAGAGGTGAATCTTACTGTGT 2640  
Qy 2641 CGATCATGAGGCTTCCCACTGTAGTCTGACAGAGGCTGATCTGAATCAAGTCCACTG 2700  
Db 2641 CGATCATGAGGCTTCCCACTGTAGTCTGACAGAGGCTGATCTGAATCAAGTCCACTG 2700  
Qy 2701 GACAGGCTGAGCTGTTTCTTTCACCTTGGCTTATGGGTAATCTCATGCTGACGAC 2760  
Db 2701 GACAGGCTGAGCTGTTTCTTTCACCTTGGCTTATGGGTAATCTCATGCTGACGAC 2760  
Qy 2761 CTGAGCCTTAAGATAAACCTGTGGAAGAAACAATGCGTGAAGCTTCTGTGCGAGGTCATG 2820  
Db 2761 CTGAGCCTTAAGATAAACCTGTGGAAGAAACAATGCGTGAAGCTTCTGTGCGAGGTCATG 2820  
Qy 2821 AGAGAACCATCTTGTATCTCCAGAGACCTGAGGTTGTAAAGTGTCACTCCACGCCCGC 2880  
Db 2821 AGAGAACCATCTTGTATCTCCAGAGACCTGAGGTTGTAAAGTGTCACTCCACGCCCGC 2880

Qy 2881 TGTGTGAGAGTCTGTCTGTGTGTATCTCGAGAGAGACACTGGAAGGCTGTGATCTC 2940  
Db 2881 TGTGTGAGAGTCTGTCTGTGTGTATCTCGAGAGAGACACTGGAAGGCTGTGATCTC 2940  
Qy 2941 ACGGACATGCGCTGGGTGACGAGTGGGTTCTCGCTGTGCGAGGAGCTGAACAAAG 3000  
Db 2941 ACGGACATGCGCTGGGTGACGAGTGGGTTCTCGCTGTGCGAGGAGCTGAACAAAG 3000  
Qy 3001 AACAGTGTCTGACGAGACTCGGGTTGAAGGCAATGTGACTGACTTGTGATTCGTGAG 3060  
Db 3001 AACAGTGTCTGACGAGACTCGGGTTGAAGGCAATGTGACTGACTTGTGATTCGTGAG 3060  
Qy 3061 GCACTTCCTGTCGCTTCTTCTGCAACCGGATGTGACCAATCTTAAACCTGTGCAAGAT 3120  
Db 3061 GCACTTCCTGTCGCTTCTTCTGCAACCGGATGTGACCAATCTTAAACCTGTGCAAGAT 3120  
Qy 3121 AACTTCAGTCCCAAGGAATGATGAAGCTGTGTGCGCTTTCCTGCTGCCAGCTTAC 3180  
Db 3121 AACTTCAGTCCCAAGGAATGATGAAGCTGTGTGCGCTTTCCTGCTGCCAGCTTAC 3180  
Qy 3181 TTACAGATTAATGGGCTGTGGAATGCGAGTACCTCTGTCAATTAAGAACTGTGAG 3240  
Db 3181 TTACAGATTAATGGGCTGTGGAATGCGAGTACCTCTGTCAATTAAGAACTGTGAG 3240  
Qy 3241 GAAGTGCAGCTACTCAAGCCCGGAGTGTAAATGACGATGTGGCATCTTTTGTATGA 3300  
Db 3241 GAAGTGCAGCTACTCAAGCCCGGAGTGTAAATGACGATGTGGCATCTTTTGTATGA 3300  
Qy 3301 GATGACCGACCAAAATAGACTTACTTTCGGGCTCCCTGAAACCGGGGCAATGGCAATG 3360  
Db 3301 GATGACCGACCAAAATAGACTTACTTTCGGGCTCCCTGAAACCGGGGCAATGGCAATG 3360  
Qy 3361 GCTTGTGTGGGGATGAACCCAGAGAGAAAGCGTGTGTGCTTGTGAGAGAC 3420  
Db 3361 GCTTGTGTGGGGATGAACCCAGAGAGAAAGCGTGTGTGCTTGTGAGAGAC 3420  
Qy 3421 TTCAAGACAGTACACGATTTTCCGAGTCTCTCTGCTGTGCAAGGCAATGTGTGATCC 3480  
Db 3421 TTCAAGACAGTACACGATTTTCCGAGTCTCTCTGCTGTGCAAGGCAATGTGTGATCC 3480  
Qy 3481 CAGAGATGACAAAGTGAAGCAGAGCTCCCGGACCCATGAGGAGGACGGAACACAA 3540  
Db 3481 CAGAGATGACAAAGTGAAGCAGAGCTCCCGGACCCATGAGGAGGACGGAACACAA 3540  
Qy 3541 CAAGTAAATATGTTGATGTGTATTTCCGAGCCTGTGTGAAACTGTGAGCTGAA 3600  
Db 3541 CAAGTAAATATGTTGATGTGTATTTCCGAGCCTGTGTGAAACTGTGAGCTGAA 3600  
Qy 3601 GGGCTTGGATCAACAGTGTGATCATGACAGAGGATGCTGCTGTCACTGAGGAGA 3660  
Db 3601 GGGCTTGGATCAACAGTGTGATCATGACAGAGGATGCTGCTGTCACTGAGGAGA 3660  
Qy 3661 GAGCTGAGCTGAAGGCTGTGTGCAACAGTGTGATGACAGAGGCTGTGTCTGTGT 3720  
Db 3661 GAGCTGAGCTGAAGGCTGTGTGCAACAGTGTGATGACAGAGGCTGTGTCTGTGT 3720  
Qy 3721 CACTGGAAGCGGCTGTGAGGCTGTGAGGCTGTGATTAAGTGTGATGACAGAGCGGT 3780  
Db 3721 CACTGGAAGCGGCTGTGAGGCTGTGAGGCTGTGATTAAGTGTGATGACAGAGCGGT 3780  
Qy 3781 GTGTCTGTGTACCTGGAAGCGGCTGTGAGGCTGTGATTAAGTGTGATGACAGAGCGGT 3840  
Db 3781 GTGTCTGTGTACCTGGAAGCGGCTGTGAGGCTGTGATTAAGTGTGATGACAGAGCGGT 3840  
Qy 3841 CACAGCGGTGTGCTGTGACCTGGAAGCGGCTGTGAGGCTGTGATTAAGTGTGATGACAG 3900  
Db 3841 CACAGCGGTGTGCTGTGACCTGGAAGCGGCTGTGAGGCTGTGATTAAGTGTGATGACAG 3900  
Qy 3901 GCTGATGACACAGCGGTGTGCTGTGATGACCTGGAAGCGGCTGTGAGGCTGTGATGACAG 3960  
Db 3901 GCTGATGACACAGCGGTGTGCTGTGATGACCTGGAAGCGGCTGTGAGGCTGTGATGACAG 3960

QY	3961	TCCAACAGTGCATGATGACCAACGCGGTGTGTCCTCTGTCACTGGAGCGCGTGGCTCGAG	4021
DB	3961	TCCAACAGTGCATGATGACCAACGCGGTGTGTCCTCTGTCACTGGAGCGCGTGGCTCGAG	4020
QY	4021	GCGCTGTGTCTTAA	4035
DB	4021	GCGCTGTGTCTTAA	4035
RESULT 2			
ADP94744			
ID	ADP94744	standard; cDNA; 4035 BP.	
AC	ADP94744;		
XX			
DT	26-FEB-2004	(first entry)	
XX			
DE	Human PYRIN-5 coding sequence.		
XX			
KW	human; PYRIN; inflammatory disorder; inappropriate apoptosis;		
KW	inflammatory bowel disease; rheumatoid arthritis; diabetes;		
KW	multiple sclerosis; Grave's disease; contact dermatitis; psoriasis;		
KW	graft rejection; asthma; allergy; chronic obstructive pulmonary disease;		
KW	glomerulonephritis; infection; Alzheimer's disease; Parkinson's disease;		
KW	anaemia; ischaemia; screening; chromosomal mapping; tissue typing;		
KW	forensic biology; pharmacogenomics; predictive medicine; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO2003089588-A2.		
XX			
PD	30-OCT-2003.		
XX			
PF	14-APR-2003; 2003WO-US011572.		
XX			
PR	17-APR-2002; 2002US-00124498.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX	(AMHP) WYETH.		
PI	Bertin J, Wang W, Blatcher M;		
XX			
DR	WPI: 2003-845527/78.		
XX			
DR	P-PSDB; ADP94745.		
XX			
PT	New nucleic acid molecules and polypeptides (e.g. PYRIN-2 or PYRIN-3)		
XX	useful for diagnosing, preventing or treating inflammation or disorders		
XX	associated with inappropriate apoptosis, in chromosomal mapping or in		
XX	pharmacogenomics.		
XX			
PS	Claim 4; SEQ ID NO 5; 199p; English.		
XX			
CC	The invention comprises the amino acid and coding sequences of human		
CC	PYRIN proteins. The DNA and protein sequences of the invention are useful		
CC	in diagnosing, preventing and treating inflammatory disorders or		
CC	disorders associated with inappropriate apoptosis, such as: inflammatory		
CC	bowel disease, rheumatoid arthritis, diabetes, graft rejection, sclerostis,		
CC	Grave's disease, contact dermatitis, psoriasis, multiple sclerosis, asthma,		
CC	allergy, chronic obstructive pulmonary disease, glomerulonephritis,		
CC	infections, Alzheimer's disease, Parkinson's disease, anaemia and		
CC	ischaemia. The DNA and protein sequences of the invention may also be		
CC	used in screening assays, chromosomal mapping, tissue typing, forensic		
CC	biology, pharmacogenomics, predictive medicine, and in monitoring of		
CC	clinical trials. The present DNA sequence encodes a PYRIN protein of the		
CC	invention.		
XX			
SQ	Sequence 4035 BP; 935 A; 1022 C; 1175 G; 903 T; 0 U; 0 Other;		
Query Match	100.0%;	Score 4035;	DB 10; Length 4035;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 4035; Conservative	0;	Mismatches	0; Indels 0; Gaps 0;

Db	1	ATGGAAGAGACDAAATCGCTAACCTTTTCCAGCTACGGGCTGGCATGATGTGTCTTAATGAG	60
QY	61	CTAGACAAGAAAGATTTACAGACATTCAGGAATTAATAAGAAATCTTCAGAAATCG	120
Db	61	CTAGACAAGAAAGATTTACAGACATTCAGGAATTAATAAGAAAGAAATCTTCAGAAATCG	120
QY	121	ACCACATGCTCTATTCCACAGTTGGAAATCGAATAATGCAACGTGGAAATGTCTGGCACTC	180
Db	121	ACCACATGCTCTATTCCACAGTTGGAAATCGAATAATGCAACGTGGAAATGTCTGGCACTC	180
QY	181	CTCTTGACATGATATTATAGAGCATCGCTGGCTGGGCTACGTCCATTAGCATTTGAA	240
Db	181	CTCTTGACATGATATTATAGAGCATCGCTGGCTGGGCTACGTCCATTAGCATTTGAA	240
QY	241	AACATGAACCTGCGNAACCCCTCTCGAGAAAGCACGGATGACATGAATAAAAAATTTCAAA	300
Db	241	AACATGAACCTGCGNAACCCCTCTCGAGAAAGCACGGATGACATGAATAAAAAATTTCAAA	300
QY	301	GCTATGGAACAAGAGGTGCCACACGACAGACAGACAGAAACAAGAAATTTCAACACT	360
Db	301	GCTATGGAACAAGAGGTGCCACACGACAGACAGACAGAAACAAGAAATTTCAACACT	360
QY	361	ATGGAACAAGAGGTGCCACACGACAGACAGACAGAAACAAGAAACAAGAAATTTCAACACT	420
Db	361	ATGGAACAAGAGGTGCCACACGACAGACAGACAGAAACAAGAAACAAGAAATTTCAACACT	420
QY	421	TGGGACTCAAGAGTCAAGTATACCAAAATTCCTGTGAGAGAGAGATGTACGTCTAGT	480
Db	421	TGGGACTCAAGAGTCAAGTATACCAAAATTCCTGTGAGAGAGAGATGTACGTCTAGT	480
QY	481	TTTGAACCACTGCTGCTGACTGGCCCGGAATGCAAACTGTGGTGTCTTTTGATTCA	540
Db	481	TTTGAACCACTGCTGCTGACTGGCCCGGAATGCAAACTGTGGTGTCTTTTGATTCA	540
QY	541	GACCGGTGGGGCTTCCGGCTTCGCAACGATGTCTTGCAACGAAATGACGAAATTTGGGAAA	600
Db	541	GACCGGTGGGGCTTCCGGCTTCGCAACGATGTCTTGCAACGAAATGACGAAATTTGGGAAA	600
QY	601	TCGGCTTAGCCAGAAAGATCGTCTGTGCTGGGCGCAAGGTGACCTTACAGAGGAATG	660
Db	601	TCGGCTTAGCCAGAAAGATCGTCTGTGCTGGGCGCAAGGTGACCTTACAGAGGAATG	660
QY	661	TTCTCTCAACGTCTTCTTCTCCCGTTAAGAGATGCAACGCGAAGAAAGAGACAGTCTC	720
Db	661	TTCTCTCAACGTCTTCTTCTCCCGTTAAGAGATGCAACGCGAAGAAAGAGACAGTCTC	720
QY	721	ACAGAGTTCATCTCCAGGAGTGGCCAGACTTCCAGAGTCCGGGTGACGGAGATCATGTCC	780
Db	721	ACAGAGTTCATCTCCAGGAGTGGCCAGACTTCCAGAGTCCGGGTGACGGAGATCATGTCC	780
QY	781	CGACCAGAAAGGCTGTGTTTCATGATTAACGCTTTCATGATTAACGCTGTCTCTCAAC	840
Db	781	CGACCAGAAAGGCTGTGTTTCATGATTAACGCTTTCATGATTAACGCTGTCTCTCAAC	840
QY	841	AATGACAAAGACTCTGCAAAAGACTGGGCTGAGAAGACGCTCCGTTCAACCTCATACGC	900
Db	841	AATGACAAAGACTCTGCAAAAGACTGGGCTGAGAAGACGCTCCGTTCAACCTCATACGC	900
QY	901	AGTCTGCTGAGAAAGTCTGTCTCCCTGAGTCTTTCTGATGTCAACCTGACAGAGCTG	960
Db	901	AGTCTGCTGAGAAAGTCTGTCTCCCTGAGTCTTTCTGATGTCAACCTGACAGAGCTG	960
QY	961	GGCACAGAAAGTCTCAAGTCAAGATGTGTCTCCCGTTACCTGTTAAGTGAAGGAATC	1020
Db	961	GGCACAGAAAGTCTCAAGTCAAGATGTGTCTCCCGTTACCTGTTAAGTGAAGGAATC	1020
QY	1021	TCGGGGGAAACAAGATTCATTCTGCTCTTGAGCGCGGAGATTTGTTGACATCAAGAACA	1080
Db	1021	TCGGGGGAAACAAGATTCATTCTGCTCTTGAGCGCGGAGATTTGTTGACATCAAGAACA	1080
QY	1081	CAAGGTTTCGTGCATCATGAACAACCTGTGCTGACACAGTCCAGAGTCCCGCC	1140
Db	1081	CAAGGTTTCGTGCATCATGAACAACCTGTGCTGACACAGTCCAGAGTCCCGCC	1140

1081 CAAGGGTTCGTGCGATCATGAACAACGCTGAGCTGCTGACCAAGTGCAGGAGCCGCC 1140  
QY 1141 GTGGGCTCTCTCATCTGCGTGGCCCTGCGAGCTGCAAGAGTGTGGGGGAGAGCGTCCG 1200  
Db 1141 GTGGGCTCTCTCATCTGCGTGGCCCTGCGAGCTGCAAGAGTGTGGGGGAGAGCGTCCG 1200  
QY 1201 CCTTCACCAACGCTCAAGGCTGACGCGCTTTGTGTTCATGAGCTCAACCCCT 1260  
Db 1201 CCTTCACCAACGCTCAAGGCTGACGCGCTTTGTGTTCATGAGCTCAACCCCT 1260  
QY 1261 CGAGCGTGGTCCGCGCTGTCTCAATCTGAGAGAAAGTGTCTGTAAGCGCTTTCG 1320  
Db 1261 CGAGCGTGGTCCGCGCTGTCTCAATCTGAGAGAAAGTGTCTGTAAGCGCTTTCG 1320  
QY 1321 CTTATGCTGTGAGAGAGTGTGGAATAGGAAGTCAAGTGTGATGTGATGACACTCAG 1380  
Db 1321 CTTATGCTGTGAGAGAGTGTGGAATAGGAAGTCAAGTGTGATGTGATGACACTCAG 1380  
QY 1381 GTTCAAGAGCTCGGGAGTCTGAGCTCGGTGCTGTTCACATGAACATCCCTTCACA 1440  
Db 1381 GTTCAAGAGCTCGGGAGTCTGAGCTCGGTGCTGTTCACATGAACATCCCTTCACA 1440  
QY 1441 GACAGCACTGTGAGAGTACTACACTTCTTCACTCAGTCTCCAGGACTTGTGTC 1500  
Db 1441 GACAGCACTGTGAGAGTACTACACTTCTTCACTCAGTCTCCAGGACTTGTGTC 1500  
QY 1501 GCGTGTACTAGCTGTAGAGGCTGGAATGAGCCAGCTCTGCGCTCTGTACGTT 1560  
Db 1501 GCGTGTACTAGCTGTAGAGGCTGGAATGAGCCAGCTCTGCGCTCTGTACGTT 1560  
QY 1561 GAGAGACAAAGGCTCATGAGCTTAAACAGGCGCTTCATATCCACTCGCTTGG 1620  
Db 1561 GAGAGACAAAGGCTCATGAGCTTAAACAGGCGCTTCATATCCACTCGCTTGG 1620  
QY 1621 ATGAAGCGTTCTGTGTGGCTCTGAGCGAAGAGTAAAGAGGCACTGAGAGTCTG 1680  
Db 1621 ATGAAGCGTTCTGTGTGGCTCTGAGCGAAGAGTAAAGAGGCACTGAGAGTCTG 1680  
QY 1681 CTGGGCTGTCCCGTTCCCTGCGGGGTGAAGCAAGAGCTTCTGACTGGGTCTCTGTTG 1740  
Db 1681 CTGGGCTGTCCCGTTCCCTGCGGGGTGAAGCAAGAGCTTCTGACTGGGTCTCTGTTG 1740  
QY 1741 GGTGACGAGCTTAATGACCAACCCAGAGAGACCCCTGAGAGCTTCCACTGCTTTTC 1800  
Db 1741 GGTGACGAGCTTAATGACCAACCCAGAGAGACCCCTGAGAGCTTCCACTGCTTTTC 1800  
QY 1801 GAGACTCAAGACAAAGAGTGTGTGCTTGGATTAAACAGCTTCCAAAGAGTGTGCTT 1860  
Db 1801 GAGACTCAAGACAAAGAGTGTGTGCTTGGATTAAACAGCTTCCAAAGAGTGTGCTT 1860  
QY 1861 CCGATTAAACAGAACTTGAATGAGATTTCTCTCTGCTCAGACATGTCGAT 1920  
Db 1861 CCGATTAAACAGAACTTGAATGAGATTTCTCTCTGCTCAGACATGTCGAT 1920  
QY 1921 TTGGGAGAAATTCGGGTGAGTGTCAAGAGGATCTTCCAAAGAGAGTCCCTGAGGA 1980  
Db 1921 TTGGGAGAAATTCGGGTGAGTGTCAAGAGGATCTTCCAAAGAGAGTCCCTGAGGA 1980  
QY 1981 TGTCTGTGTGCTCTATGATGAGCGGATAGAGCCCTCATTTGAGAGAGTGTGGAAGT 2040  
Db 1981 TGTCTGTGTGCTCTATGATGAGCGGATAGAGCCCTCATTTGAGAGAGTGTGGAAGT 2040  
QY 2041 TTCTGCTTCATGCTTGGACCCACCACTGCGGAGCTGACCTGGGAGAGCATC 2100  
Db 2041 TTCTGCTTCATGCTTGGACCCACCACTGCGGAGCTGACCTGGGAGAGCATC 2100  
QY 2101 CTGAGAGCGGGCCATGAAAGCCGTGTGCGAAGCTGAGGACTCCACTGCAAGATA 2160  
Db 2101 CTGAGAGCGGGCCATGAAAGCCGTGTGCGAAGCTGAGGACTCCACTGCAAGATA 2160  
QY 2161 CAGACCTGATGTTAGAAATGACAGATTACCTCTGTGTGACACCTCTGGAGATC 2220  
Db 2161 CAGACCTGATGTTAGAAATGACAGATTACCTCTGTGTGACACCTCTGGAGATC 2220

QY 2221 GTATGAGCCAAACCTTAACCTTAAGATCCCTCAACTTGGAGAGGACCCACTGGAAGAG 2280  
Db 2221 GTATGAGCCAAACCTTAACCTTAAGATCCCTCAACTTGGAGAGGACCCACTGGAAGAG 2280  
QY 2281 GATGTAAAGATGAGCTGTGAAGCTTAAACCAACCAAAATGTTGTGAGCTTGAAG 2340  
Db 2281 GATGTAAAGATGAGCTGTGAAGCTTAAACCAACCAAAATGTTGTGAGCTTGAAG 2340  
QY 2341 CTGATGTGCTGTGATGACCCATGCTGTAACTGAAAGATCTCCAAATCTTACGAC 2400  
Db 2341 CTGATGTGCTGTGATGACCCATGCTGTAACTGAAAGATCTCCAAATCTTACGAC 2400  
QY 2401 TCCCCAGCTGAAATCTCTGAGCTGTGAGAGAAACAGGTGACAGACAGGAGATATG 2460  
Db 2401 TCCCCAGCTGAAATCTCTGAGCTGTGAGAGAAACAGGTGACAGACAGGAGATATG 2460  
QY 2461 CCTCTCAGTATGCTTGAAGTCTCCAGTGCAGGCTGAGAGCTGATCTGAGAGAC 2520  
Db 2461 CCTCTCAGTATGCTTGAAGTCTCCAGTGCAGGCTGAGAGCTGATCTGAGAGAC 2520  
QY 2521 TGTGCACTACAGCCAGGCTTGCAGAGTCTGAGCTCAGCCCTGTGACAAACCGAGC 2580  
Db 2521 TGTGCACTACAGCCAGGCTTGCAGAGTCTGAGCTCAGCCCTGTGACAAACCGAGC 2580  
QY 2581 TTGACACACTGTGCTTATCCAAACAGCTGCGGAAACGAAGGTGTAATCTACTGT 2640  
Db 2581 TTGACACACTGTGCTTATCCAAACAGCTGCGGAAACGAAGGTGTAATCTACTGT 2640  
QY 2641 CGATCATGAGGCTTCCCACTGTGTGAGAGAGGCTGAGATGCTGAATCAGTGCACCTG 2700  
Db 2641 CGATCATGAGGCTTCCCACTGTGTGAGAGAGGCTGAGATGCTGAATCAGTGCACCTG 2700  
QY 2701 GACAGGCTGTGCTGTGTTTCTTGCACCTTGCTTAATGAGTACTCATGCTGACGAC 2760  
Db 2701 GACAGGCTGTGCTGTGTTTCTTGCACCTTGCTTAATGAGTACTCATGCTGACGAC 2760  
QY 2761 CTGAGCTTATGATTAACCTGTGTGAAGACAAATGCGTGAAGCTTCTGTGAGTCAATG 2820  
Db 2761 CTGAGCTTATGATTAACCTGTGTGAAGACAAATGCGTGAAGCTTCTGTGAGTCAATG 2820  
QY 2821 AGAGAACATCTGTGATCTCCAGGACCTGAGATGTGAAGTGTCACTCAACGCGCG 2880  
Db 2821 AGAGAACATCTGTGATCTCCAGGACCTGAGATGTGAAGTGTCACTCAACGCGCG 2880  
QY 2881 TGTGTGAGAGTCTGTCTGTGTGATCTGAGAGACAGACACTGAAAGCTGTGATCTC 2940  
Db 2881 TGTGTGAGAGTCTGTCTGTGTGATCTGAGAGACAGACACTGAAAGCTGTGATCTC 2940  
QY 2941 ACGACAAATGCCCTGGTGAAGGCTGTGCGTGTGCGTGTGCGAGGAGCTGAAGCAAAAG 3000  
Db 2941 ACGACAAATGCCCTGGTGAAGGCTGTGCGTGTGCGTGTGCGAGGAGCTGAAGCAAAAG 3000  
QY 3001 AACAGTGTGTGACGAGCTGGGTTGAAGGATGTGAAGCTTGTGATTTGCTGTAG 3060  
Db 3001 AACAGTGTGTGACGAGCTGGGTTGAAGGATGTGAAGCTTGTGATTTGCTGTAG 3060  
QY 3061 GCACTCTCTTGGCCCTTCTGCAACCGGATCTGACCACTGTAACCTGTGTGAGAT 3120  
Db 3061 GCACTCTCTTGGCCCTTCTGCAACCGGATCTGACCACTGTAACCTGTGTGAGAT 3120  
QY 3121 AACTTCACTCCAAAGAAATGATGAAGTGTGCTTGGCTTGTCCACGCTTAAC 3180  
Db 3121 AACTTCACTCCAAAGAAATGATGAAGTGTGCTTGGCTTGTCCACGCTTAAC 3180  
QY 3181 TTACAGATTAATTTGGCTGTGAAATGAGAGTGTGCAATTAAGAACTGTGTGAG 3240  
Db 3181 TTACAGATTAATTTGGCTGTGAAATGAGAGTGTGCAATTAAGAACTGTGTGAG 3240  
QY 3241 GAAAGTGAAGTACTCAAGCCCGAGTGTAAATGAGCTGATGAGCTTCTTTGATGA 3300  
Db 3241 GAAAGTGAAGTACTCAAGCCCGAGTGTAAATGAGCTGATGAGCTTCTTTGATGA 3300



```

QY 3301 GATGACCGACACAAAATAGACTTACTTCCGGCTCCCTGAAAGCCGGGCAATGGCCATGT 3360
Db 3301 GATGACCGACACAAAATAGACTTACTTCCGGCTCCCTGAAAGCCGGGCAATGGCCATGT 3360
QY 3361 GCCTTGCTGTGGGGATGAACCCAGAGCAAGAAAGCTGTGTCCCTTCTGGCTGGAGAC 3420
Db 3361 GCCTTGCTGTGGGGATGAACCCAGAGCAAGAAAGCTGTGTCCCTTCTGGCTGGAGAC 3420
QY 3421 TTCAAGACAGTACAGATTTGGCCAAAGTCTCTCTGGCTGGCCAGGCAAAATGGTAGTCC 3480
Db 3421 TTCAAGACAGTACAGATTTGGCCAAAGTCTCTCTGGCTGGCCAGGCAAAATGGTAGTCC 3480
QY 3481 CAGAGAGTTGACAAACGTGAGAGAGCTCCCGCAACCCATGAGGAGGACGGAACACAAA 3540
Db 3481 CAGAGAGTTGACAAACGTGAGAGAGCTCCCGCAACCCATGAGGAGGACGGAACACAAA 3540
QY 3541 CAGATTAATAATGTGATGTGATATTCGAGAGCTGTGCTGAAACTGTGAGCTGAA 3600
Db 3541 CAGATTAATAATGTGATGTGATATTCGAGAGCTGTGCTGAAACTGTGAGCTGAA 3600
QY 3601 GGGCTTGATCCAAAGTGTGATGATGACCAAGGAGGATAGGCTGTGCTAGAGGAGA 3660
Db 3601 GGGCTTGATCCAAAGTGTGATGATGACCAAGGAGGATAGGCTGTGCTAGAGGAGA 3660
QY 3661 GAGCTGAGCTGAGGGGCTGTGTCTCAACAGTGTGATGACCAAGCGGTGTCTTGT 3720
Db 3661 GAGCTGAGCTGAGGGGCTGTGTCTCAACAGTGTGATGACCAAGCGGTGTCTTGT 3720
QY 3721 CACTGGAGCGGCTGTGGCTCTAGGGGCTGTGTCTTAAACAGTGTGATGACCAAGCGGT 3780
Db 3721 CACTGGAGCGGCTGTGGCTCTAGGGGCTGTGTCTTAAACAGTGTGATGACCAAGCGGT 3780
QY 3781 GGTCTCTGATCACTGGAGCGGCTGTGGCTGTGTCTTAAACAGTGTGATGAC 3840
Db 3781 GGTCTCTGATCACTGGAGCGGCTGTGGCTGTGTCTTAAACAGTGTGATGAC 3840
QY 3841 CACAAGCGGTGTGGCTGTGTCTTAAACAGTGTGATGACCAAGCGGTGTGTCTTAAACAGT 3900
Db 3841 CACAAGCGGTGTGGCTGTGTCTTAAACAGTGTGATGACCAAGCGGTGTGTCTTAAACAGT 3900
QY 3901 GCTGATGACCAAGCGGTGTGTCTTGTGATGACCAAGCGGTGTGTCTTGTG 3960
Db 3901 GCTGATGACCAAGCGGTGTGTCTTGTGATGACCAAGCGGTGTGTCTTGTG 3960
QY 3961 TCCAAAGTGTGATGACCAAGCGGTGTGTCTTGTGATGACCAAGCGGTGTGTG 4020
Db 3961 TCCAAAGTGTGATGACCAAGCGGTGTGTCTTGTGATGACCAAGCGGTGTGTG 4020
QY 4021 GGGCTGTGTCTTAA 4035
Db 4021 GGGCTGTGTCTTAA 4035

RESULT 3
AAD41224
ID AAD41224 standard; cDNA; 3489 BP.
XX AC AAD41224;
XX DT 30-OCT-2002 (first entry)
XX DE Human EMBRY-1 cDNA.
XX KW Human; embryogenesis associated protein; AIDS; reproductive disorder;
XX KW infertility; endometriosis; endometrial tumour; inflammatory disorder;
XX KW autoimmune disorder; acquired immune deficiency syndrome; transgenic;
XX KW ovarian tumour; contact dermatitis; placenta disorder; preclampsia;
XX KW EMBRY-1; allergy; gene therapy; gene; ss.
OS Homo sapiens.
XX Key Location/Qualifiers
FH CDS 1..3489
FT

```

```

FT /*tag= a
FT /product= "EMBRY-1 protein"
XX
XX WO200248362-A2.
XX
XX 20-JUN-2002.
XX
XX 14-NOV-2001; 2001WO-US043956.
XX
XX 15-NOV-2000; 2000US-0249407P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Rankumar J, Arvizu C;
XX
XX WPI: 2002-537629/57.
XX
XX P-PSDB; AAB25053.
XX
XX New polypeptides of human embryogenesis associated proteins for screening
XX modulators useful for treating or preventing disorders e.g.
XX endometriosis, infertility, allergy, preclampsia.
XX
XX Claim 58; Page 95-96; 97p; English.
XX
XX The invention relates to human embryogenesis associated proteins (EMBRY)
XX and nucleic acid molecules encoding such proteins. EMBRY sequences are
XX useful for screening modulators useful for treating or preventing
XX disorders associated with abnormal expression of EMBRY. The disorders
XX treated include reproductive disorders such as infertility,
XX endometriosis, endometrial or ovarian tumour; autoimmune/inflammatory
XX disorder such as acquired immune deficiency syndrome (AIDS), allergies,
XX contact dermatitis; disorders of the placenta such as preclampsia,
XX abruptio placentae etc. Sequences of the invention are also useful for
XX analysing a proteome of a tissue or a cell type. EMBRY proteins are
XX useful as immunogens for preparing antibodies. Polynucleotides of the
XX invention are useful for creating knockin humanised animals or transgenic
XX animals to model human diseases. They are also used in gene therapy. The
XX present sequence is human EMBRY-1 cDNA
XX
XX Sequence 3489 BP; 854 A; 881 C; 961 G; 793 T; 0 U; 0 Other;
XX
XX Query Match 72.3%; Score 2917; DB 6; Length 3489;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 3017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 290 AAATTTCAAGCTATGGAACAAGAGTCCACAGAGAGAGACAGAGAAAGAA 349
Db 452 AAATTTCAAGCTATGGAACAAGAGTCCACAGAGAGAGACAGAGAAAGAA 511
QY 350 TTTCAAGCTATGGAACAAGAGTCCACAGAGAGAGACAGAGAAAGAAAGCATG 409
Db 512 TTTCAAGCTATGGAACAAGAGTCCACAGAGAGAGACAGAGAAAGAAAGCATG 571
QY 410 GAGGTGACACATGGAGCTACAGAGTCAAGTATGACCAATTCCTGAGAGAGCATG 469
Db 572 GAGGTGACACATGGAGCTACAGAGTCAAGTATGACCAATTCCTGAGAGAGCATG 631
QY 470 TAGCTGATGTTTGAACAACCTGCTGACATGGCCGGAATGCAACGTTGGCTGGT 529
Db 632 TAGCTGATGTTTGAACAACCTGCTGACATGGCCGGAATGCAACGTTGGCTGGT 691
QY 530 CTTTGAATGACACCGGTGGGGCTTCCGGCTCGACAGTGTCTTCTGACAGAAAGTCAG 589
Db 692 CTTTGAATGACACCGGTGGGGCTTCCGGCTCGACAGTGTCTTCTGACAGAAAGTCAG 751
QY 590 GAATTTGGAATTCGGCTCTTACCAAGAGATGCTGTCTGTGGGCAAGTGTGACTCT 649
Db 752 GAATTTGGAATTCGGCTCTTACCAAGAGATGCTGTCTGTGGGCAAGTGTGACTCT 811
QY 650 ACCAGGAATGTTCTCTAGCTTCTTCTTCTCCCGTTAGAGAGTCAACGGAAGAG 709
Db 812 ACCAGGAATGTTCTCTAGCTTCTTCTTCTCCCGTTAGAGAGTCAACGGAAGAG 871

```



QY 710 AGAGCAGTGTCA CAGAGTTCA TCTCCAGGAGAGTGCCACAGCTCCAGGCTCCGGTACCG 769  
| | | | |  
Db 872 AAGAGAGTGTCA CAGAGTTCA TCTCCAGGAGAGTGCCACAGCTCCAGGCTCCGGTACCG 931  
| | | | |  
QY 770 AATATATGTCCGCA CAGAAAGGCTGTGTTCATCATTTGACGGTTTGCATGACCTTGAGCT 829  
| | | | |  
Db 932 AATATATGTCCGCA CAGAAAGGCTGTGTTCATCATTTGACGGTTTGCATGACCTTGAGCT 991  
| | | | |  
QY 830 CTGTCTCTCAAA TGA CAAAGAGCTGTGCAAGAGCTGGGCTGAGAAAGCAGCTCCGTTCA 889  
| | | | |  
Db 992 CTGTCTCTCAAA TGA CAAAGAGCTGTGCAAGAGCTGGGCTGAGAAAGCAGCTCCGTTCA 1051  
| | | | |  
QY 890 CCTCATACGAGTGTGCTGAGAGAGTCTGCTCCCTGAGTCTTCCGTTACGATGTCACCG 949  
| | | | |  
Db 1052 CCTCATACGAGTGTGCTGAGAGAGTCTGCTCCCTGAGTCTTCCGTTACGATGTCACCG 1111  
| | | | |  
QY 950 TCAGAGACGTGGGCA CAGAGAGCTCAAGTCAAGAGTCTGCTCCCTGATCTGTTAG 1009  
| | | | |  
Db 1112 TCAGAGACGTGGGCA CAGAGAGCTCAAGTCAAGAGTCTGCTCCCTGATCTGTTAG 1171  
| | | | |  
QY 1010 TTAGAGGAATCTCCGGGAA CAAAGATCACTTGCTCTTGAAGCGGGATTTGGTAC 1069  
| | | | |  
Db 1172 TTAGAGGAATCTCCGGGAA CAAAGATCACTTGCTCTTGAAGCGGGATTTGGTAC 1231  
| | | | |  
QY 1070 ATCAGAGACAGAGGGTTGCGGCGATCATGAA CAAAGCTGAGCTGCTGACAGTGC 1129  
| | | | |  
Db 1232 ATCAGAGACAGAGGGTTGCGGCGATCATGAA CAAAGCTGAGCTGCTGACAGTGC 1291  
| | | | |  
QY 1130 AGGTGCGCGCGTGGGCTCTCATCTGCGTGGCCCTGAGCTGACAGAGCTGTGGGG 1189  
| | | | |  
Db 1292 AGGTGCGCGCGTGGGCTCTCATCTGCGTGGCCCTGAGCTGACAGAGCTGTGGGG 1351  
| | | | |  
QY 1190 AAGAGTGCCTCCCTTCA CAAAGCTCA CAGAGCTGCA CCGCGCTTTTGTGTTTCA TC 1249  
| | | | |  
Db 1352 AAGAGTGCCTCCCTTCA CAAAGCTCA CAGAGCTGCA CCGCGCTTTTGTGTTTCA TC 1411  
| | | | |  
QY 1250 AGCTCA CCGCTGAGAGGCTGTGCTCCGCGCTGCTCAATCTGAGAGAAAGTTGTCMA 1309  
| | | | |  
Db 1412 AGCTCA CCGCTGAGAGGCTGTGCTCCGCGCTGCTCAATCTGAGAGAAAGTTGTCMA 1471  
| | | | |  
QY 1310 AGCGCTTCTGCGGTATGCTGTGAGAGAGTGTGAAATAGAAAGTCAAGTTGATGTC 1369  
| | | | |  
Db 1472 AGCGCTTCTGCGGTATGCTGTGAGAGAGTGTGAAATAGAAAGTCAAGTTGATGTC 1531  
| | | | |  
QY 1370 ACGACCTCATGTTCAAGAGCTCGGGAGTCTGAGCTCCGTGCTGTGTTCA CATTGACA 1429  
| | | | |  
Db 1532 ACGACCTCATGTTCAAGAGCTCGGGAGTCTGAGCTCCGTGCTGTGTTCA CATTGACA 1591  
| | | | |  
QY 1430 TCCCTTCTCCGAGACAGCCCATGTGAGAGATCA CACTTTTCCACTTCACTTCCAG 1489  
| | | | |  
Db 1592 TCCCTTCTCCGAGACAGCCCATGTGAGAGATCA CACTTTTCCACTTCACTTCCAG 1651  
| | | | |  
QY 1490 ACTTCTGTGCGCGCTTGTACTAGTGTGAGAGGCTGTGAAATTCAGAGCAGCTTCTGCG 1549  
| | | | |  
Db 1652 ACTTCTGTGCGCGCTTGTACTAGTGTGAGAGGCTGTGAAATTCAGAGCAGCTTCTGCG 1711  
| | | | |  
QY 1550 CTCTGTACGTTGAGAA CAAAGAGTTCATGAGAGCTTAAACAGAGGCTTCCATATCC 1609  
| | | | |  
Db 1712 CTCTGTACGTTGAGAA CAAAGAGTTCATGAGAGCTTAAACAGAGGCTTCCATATCC 1771  
| | | | |  
QY 1610 ACTGCGTTGAGTGAAGCGTTTCTGTTGCGCTGAGCGAAGAGCTTAAAGAGGCCAC 1669  
| | | | |  
Db 1772 ACTGCGTTGAGTGAAGCGTTTCTGTTGCGCTGAGCGAAGAGCTTAAAGAGGCCAC 1831  
| | | | |  
QY 1670 TGGAGGCTCTGCTGGGCTGTCCCGTTCCCTGGGGGTGAGAGCAAGCTTCTGCACTGG 1729  
| | | | |  
Db 1832 TGGAGGCTCTGCTGGGCTGTCCCGTTCCCTGGGGGTGAGAGCAAGCTTCTGCACTGG 1891  
| | | | |  
QY 1730 TCTCTCTGTGGGTGAGAGCTTATGCA CCA CCGCAGAGAGACACCTTGAAGCGCTTCC 1789  
| | | | |  
Db 1892 TCTCTCTGTGGGTGAGAGCTTATGCA CCA CCGCAGAGAGACACCTTGAAGCGCTTCC 1951  
| | | | |  
QY 1790 ACTGTCTTTTGAAGCTCAAGAGCTTGTGCTGCTTGGCATTTAAACAGCTTCCAG 1849  
| | | | |

Db 1952 ACTGTCTTTTGAAGCTCAAGAGAGTGTGTTGCTGTCGATTTAAACAGCTTCCAG 2011  
| | | | |  
QY 1850 AAGTGTGCTTCCGATTA CCAAGACTTGAATGATGATCTTCTTCTGCTCCAGC 1909  
| | | | |  
Db 2012 AAGTGTGCTTCCGATTA CCAAGACTTGAATGATGATCTTCTTCTGCTCCAGC 2071  
| | | | |  
QY 1910 ACTGTCCGATTTTGGGAAAAATTCGGGTGAGATGTCMAAGGATCTTCCCAAGATGAGT 1969  
| | | | |  
Db 2072 ACTGTCCGATTTTGGGAAAAATTCGGGTGAGATGTCMAAGGATCTTCCCAAGATGAGT 2131  
| | | | |  
QY 1970 CCGTGAAGCATGTCTGTGCTCCCTGATGAGATCGGGATTAAGACCTCATTTGAGAGC 2029  
| | | | |  
Db 2132 CCGTGAAGCATGTCTGTGCTCCCTGATGAGATCGGGATTAAGACCTCATTTGAGAGC 2191  
| | | | |  
QY 2030 AGTGGAAAGATTTGCTGCTCATGCTTGACCCACCA CACTGCGGAGCTGACCTGG 2089  
| | | | |  
Db 2192 AGTGGAAAGATTTGCTGCTCATGCTTGACCCACCA CACTGCGGAGCTGACCTGG 2251  
| | | | |  
QY 2090 GAGCAGCATCTGACAGAGCGGGCATGAAAGACCTGTGTGTCMAAGCTGAGGATCCCA 2149  
| | | | |  
Db 2252 GAGCAGCATCTGACAGAGCGGGCATGAAAGACCTGTGTGTCMAAGCTGAGGATCCCA 2311  
| | | | |  
QY 2150 CCTGCAAGATACAGACCCGATGTTAAGAAATGCA CAGATTA CCGCTGGTGTGACAGC 2209  
| | | | |  
Db 2312 CCTGCAAGATACAGACCCGATGTTAAGAAATGCA CAGATTA CCGCTGGTGTGACAGC 2371  
| | | | |  
QY 2210 TCTGAGATATGTCATGAGCCAA CCGTAA CTTAAGTCCCTCAACTTGGAGGACCCAC 2269  
| | | | |  
Db 2372 TCTGAGATATGTCATGAGCCAA CCGTAA CTTAAGTCCCTCAACTTGGAGGACCCAC 2431  
| | | | |  
QY 2270 TGAAGAAAGAGATGTAAGATGCGGTGAAAGCTTAAACACCAAAATTTGTTGG 2329  
| | | | |  
Db 2432 TGAAGAAAGAGATGTAAGATGCGGTGAAAGCTTAAACACCAAAATTTGTTGG 2491  
| | | | |  
QY 2330 AGTCTTTGAGGCTGAGTGTGCTGTGATGACCCATGAGCTGTTACCTGAAGATCTCCCAA 2389  
| | | | |  
Db 2492 AGTCTTTGAGGCTGAGTGTGCTGTGATGACCCATGAGCTGTTACCTGAAGATCTCCCAA 2551  
| | | | |  
QY 2390 TCCTTACGACCTCCCA CAGCGCTGAAATCTCTGAGCTGTGGAGAA CAAAGTGA CAGAC 2449  
| | | | |  
Db 2552 TCCTTACGACCTCCCA CAGCGCTGAAATCTCTGAGCTGTGGAGAA CAAAGTGA CAGAC 2611  
| | | | |  
QY 2450 AGGAGTAAATGCTTCAAGTATGCTTGAAGTCTCCAGTGCCTGCAAGAGCTGA 2509  
| | | | |  
Db 2612 AGGAGTAAATGCTTCAAGTATGCTTGAAGTCTCCAGTGCCTGCAAGAGCTGA 2671  
| | | | |  
QY 2510 TACTGAGAGACGTGTGCAATCA CAGCAGAGGTGCAAGTGTGCTCAGGCCCTGCA 2569  
| | | | |  
Db 2672 TACTGAGAGACGTGTGCAATCA CAGCAGAGGTGCAAGTGTGCTCAGGCCCTGCA 2731  
| | | | |  
QY 2570 GCAACCGAGGCTTGA CACAGCTGTGCTATCCATCA CAAAGCTGTGGGAA CCAAGGTTGA 2629  
| | | | |  
Db 2732 GCAACCGAGGCTTGA CACAGCTGTGCTATCCATCA CAAAGCTGTGGGAA CCAAGGTTGA 2791  
| | | | |  
QY 2630 ATCTACTGTGATCAATGAGGCTTCCCACTGATGCTGCAAGGCTGATGCTGAATC 2689  
| | | | |  
Db 2792 ATCTACTGTGATCAATGAGGCTTCCCACTGATGCTGCAAGGCTGATGCTGAATC 2851  
| | | | |  
QY 2690 AGTGCACCTGAGACAGGCTGCTGTGTTTCTTGA CACTGCGCTTAAAGGTTACTAT 2749  
| | | | |  
Db 2852 AGTGCACCTGAGACAGGCTGCTGTGTTTCTTGA CACTGCGCTTAAAGGTTACTAT 2911  
| | | | |  
QY 2750 GGTGACGACACTGAGCTTATGATGAACCCCTGTGGAAGACAAATGCGCTGAAGCTTCTGT 2809  
| | | | |  
Db 2912 GGTGACGACACTGAGCTTATGATGAACCCCTGTGGAAGACAAATGCGCTGAAGCTTCTGT 2971  
| | | | |  
QY 2810 GCGAGGTATGAGAGAACATCTTGTATCTCAGAGACCTGAGATTGTTAAAGTCTATC 2869  
| | | | |  
Db 2972 GCGAGGTATGAGAGAACATCTTGTATCTCAGAGACCTGAGATTGTTAAAGTCTATC 3031  
| | | | |  
QY 2870 TCAACGCGCGGTGCTGAGAGTCTGCTGTGAGATCTCAGAGAGGAGACACCTGAAGA 2929  
| | | | |

Db 3032 TCACCGCGCGTGTGTGAGAGTCTGTCTGTGTGATCTCGAGAGCAGACACTGAAGA 3091  
Qy GCTTGATCTCAGCAGCAATGCCCTGGGTGA CGGTGGGGTTGTGCGCTGTGCGAGGAGC 2989  
Db GCTTGATCTCAGCAGCAATGCCCTGGGTGA CGGTGGGGTTGTGCGCTGTGCGAGGAGC 3151  
Qy TGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGCATGTGACTGACTTCTG 3049  
Db TGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGCATGTGACTGACTTCTG 3211  
Qy ATGTGTGAGGACATCTCTTGGCCCTTTCTGCAACCGGCATGTGACCACTTAAAC 3109  
Db ATGTGTGAGGACATCTCTTGGCCCTTTCTGCAACCGGCATGTGACCACTTAAAC 3271  
Qy TGTGTCAGATTAATCTTCACTGCTCCCAAGAGATGATGACTGTGTGCGCTTGGCTGTC 3169  
Db TGTGTCAGATTAATCTTCACTGCTCCCAAGAGATGATGACTGTGTGCGCTTGGCTGTC 3331  
Qy 3170 CCACGCTTAATTAATGAGATTTGGGCTGTGAAAATGCGAGTACCCTGTGCAATTAAGA 3229  
Db 3332 CCACGCTTAATTAATGAGATTTGGGCTGTGAAAATGCGAGTACCCTGTGCAATTAAGA 3391  
Qy 3230 AGCTGCTGAGAGATGCACTCAAGCCCGAGTGTATTAAGAGTGTGCGATT 3289  
Db 3392 AGCTGCTGAGAGATGCACTCAAGCCCGAGTGTATTAAGAGTGTGCGATT 3451  
Qy CTTTGTATGAGATGACCG 3308  
Db CTTTGTATGAGATGACCG 3470

RESULT 4  
ADA45218  
ID ADA45218 standard; DNA; 3926 BP.

ADA45218;  
20-NOV-2003 (first entry)

Human MATER DNA.

human ; MATER, maternal antigen that embryos require; NTP-ase;  
nucleoside triphosphate-ase; apoptosis; antiinfertility; contraceptive;  
antiflamatory; immunosuppressive; gynaecological; endometriosis;  
ovarian dysfunction; autoimmune premature ovarian failure;  
autoimmune disease; contraceptive; autoantibody; ds; gene.

Homo sapiens.

Key Location/Qualifiers  
CDS 1..3489  
FT /\*tag= a  
FT /product= "MATER"

EP1285964-A2.

26-FEB-2003.

12-JUL-2002; 2002EP-00090246.

10-AUG-2001; 2001DE-01039874.

(SCHD) SCHERING AG.

Weiss B, Lesel M, Peters-Kottig M, Beckmann G;

WPI; 2003-302814/30.

P-PSDB; ADA45219.

New nucleic acid for human maternal antigen that embryos require protein,  
useful e.g. in diagnosis and treatment of female infertility.

Claim 1; Page 9-11; 31pp; German.

XX This invention describes a novel human MATER (maternal antigen that  
CC embryos require) protein. MATER is an NTP(nucleoside triphosphate)-ase  
CC associated with apoptosis, a defect in MATER activity causes growth  
CC arrest at the 2-cell stage of fertilized eggs. mRNA transcribed from the  
CC MATER gene is found predominantly in ovary, testis and placenta. The  
CC products of the invention have antiinfertility, contraceptive,  
CC antiflamatory, immunosuppressive and gynaecological activity.  
CC Effectors of the MATER polypeptide are used to treat infertility.  
CC associated with endometriosis and also ovarian dysfunction, autoimmune  
CC premature ovarian failure, inflammation, autoimmune diseases and female  
CC infertility, and as contraceptives. Measurements of autoantibodies  
CC against MATER in body samples or MATER protein or mRNA in eggs is useful  
CC for diagnosis of female infertility, particularly by detecting mutations  
CC in the MATER gene with a DNA chip. This sequence encodes the human MATER  
CC polypeptide described in the invention.

SQ Sequence 3926 BP; 969 A; 958 C; 1061 G; 938 T; 0 U; 0 Other;

Query Match 72.3%; Score 2917; DB 8; Length 3926;

Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 290 AAATTTCACAAAGCTATGAGACAGAGAGTGCACACAGACAGACAGAGAGACAAAGAAA 349  
Db 452 AAATTTCACAAAGCTATGAGACAGAGAGTGCACACAGACAGACAGAGAGACAAAGAAA 511  
Qy 350 TTTCAAGAGTATGAGACAAAGAGTGCACAGAGAGAGACAGAGAGACAAAGAGACATG 409  
Db 512 TTTCAAGAGTATGAGACAAAGAGTGCACAGAGAGAGACAGAGAGACAAAGAGACATG 571  
Qy 410 GAGGTGACACATGAGGAGCTACAAAGAGTACAGTATGATGACCAATTCGTGAGAGAGAGATG 469  
Db 572 GAGGTGACACATGAGGAGCTACAAAGAGTACAGTATGATGACCAATTCGTGAGAGAGAGATG 631  
Qy 470 TAGCTGTAATTTGAAAAACATGCTGCTGATCTGCGCGAAATCCAAACGTTGGCTGTG 529  
Db 632 TAGCTGTAATTTGAAAAACATGCTGCTGATCTGCGCGAAATCCAAACGTTGGCTGTG 691  
Qy 530 CTTTGTATGACAGCGGGGGGCTTCGGGCTGCGACAGAGTGTCTGACAGGAAAGTCAG 589  
Db 692 CTTTGTATGACAGCGGGGGGCTTCGGGCTGCGACAGAGTGTCTGACAGGAAAGTCAG 751  
Qy 590 GAATTGGGAAATCGGCTCTAGCCAGAGATCGTGTGTGCGGCGCAAGGTGACTCT 649  
Db 752 GAATTGGGAAATCGGCTCTAGCCAGAGATCGTGTGTGCGGCGCAAGGTGACTCT 811  
Qy 650 ACCAGGGAATGTTCTCTAGCTCTTCTCTCCCGTTAGAGATGACGCGAAGAG 709  
Db 812 ACCAGGGAATGTTCTCTAGCTCTTCTCTCCCGTTAGAGATGACGCGAAGAG 871  
Qy 710 AGAGCAGGTGACAGAGTTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGGTACGG 769  
Db 872 AGAGCAGGTGACAGAGTTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGGTACGG 931  
Qy 770 AGATCATGTCCCGACAGAAAGCTGTGTTTATCATTTGACGATTTGATGACTTGAGCT 829  
Db 932 AGATCATGTCCCGACAGAAAGCTGTGTTTATCATTTGACGATTTGATGACTTGAGCT 991  
Qy 830 CTGTCTCAACATGACACAAAGCTGTGCAAGATCGGCTGAGAGCAGCTCCGTTCA 889  
Db 992 CTGTCTCAACATGACACAAAGCTGTGCAAGATCGGCTGAGAGCAGCTCCGTTCA 1051  
Qy 890 CCTTCATACGAGCTGTCTGAGAGAGTCTGCTCCCTGAGTCTCTTCTATGCTGACCG 949  
Db 1052 CCTTCATACGAGCTGTCTGAGAGAGTCTGCTCCCTGAGTCTCTTCTATGCTGACCG 1111  
Qy 950 TCAGAGAGTGGGACAGAGAGCTCAAGTCAGAGTGTGTCTCCCGTTACTGTTAG 1009  
Db 1112 TCAGAGAGTGGGACAGAGAGCTCAAGTCAGAGTGTGTCTCCCGTTACTGTTAG 1171  
Qy 1010 TTAGAGGAATCTCCGGGAAACAAAGATCACTTGTCTTGAAGCGGGAGTTGTAGC 1069  
[[[

Db 1172 TTAGAGAAATCTCCGGGAAACAAAGAAATCATTGTGCTCTTGAGCGCGGGATTGGTAGAC 1231  
Qy 1070 ATCAGAAACACAAGGTTGCGTGATATGAACAACCTGAGCTGCTCAGCACTGCC 1129  
Db 1232 ATCAGAAACACAAGGTTGCGTGATATGAACAACCTGAGCTGCTCAGCACTGCC 1291  
Qy 1130 AGGTGCGCGCGTGGGCTCTCATCTGCGTGACCCCTGAGCTGAGAGCGTGGGGG 1189  
Db 1292 AGGTGCGCGCGTGGGCTCTCATCTGCGTGACCCCTGAGCTGAGAGCGTGGGGG 1351  
Qy 1190 AGAGCGTGCCTCCCTTCAACCAACGCTCAGAGCTGACCGCTGTTTGTTCATC 1249  
Db 1352 AGAGCGTGCCTCCCTTCAACCAACGCTCAGAGCTGACCGCTGTTTGTTCATC 1411  
Qy 1250 AGCTCACTCTGAGGCGTGGTCCGCGCTGTCTCATCTGAGAGAAAGTTTCTGA 1309  
Db 1412 AGCTCACTCTGAGGCGTGGTCCGCGCTGTCTCATCTGAGAGAAAGTTTCTGA 1471  
Qy 1310 AGCGCTTCCGCTATGCTGTGAGAGAGTGGAAATGAAAGTCAATGTTGATGTTG 1369  
Db 1472 AGCGCTTCCGCTATGCTGTGAGAGAGTGGAAATGAAAGTCAATGTTGATGTTG 1531  
Qy 1370 AGACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCCGCTCTGTTTCAATGAACA 1429  
Db 1532 AGACCTCATGTTCAAGGACTCGGGAGTCTGAGCTCCGCTCTGTTTCAATGAACA 1591  
Qy 1430 TCTCTTCTCCAGACAGCCACTGTGAGAGTACTACCTTCTTCCACTCTGACTTCCAG 1489  
Db 1592 TCTCTTCTCCAGACAGCCACTGTGAGAGTACTACCTTCTTCCACTCTGACTTCCAG 1651  
Qy 1490 ACTCTGTGCGCGCTTGTACTAGTGTGAGAGGCTGGAATCCAGCGACGCTCTGCG 1549  
Db 1652 ACTCTGTGCGCGCTTGTACTAGTGTGAGAGGCTGGAATCCAGCGACGCTCTGCG 1711  
Qy 1550 CTCTGTACTGTGAAGAACAAAGAGTCCATGAGCTTAAACAGGAGCTTCCATATCC 1609  
Db 1712 CTCTGTACTGTGAAGAACAAAGAGTCCATGAGCTTAAACAGGAGCTTCCATATCC 1771  
Qy 1610 ACTCGCTTGTGATGAAGCGCTTCTGTTGGCCTGTGAGCGAAGAGTAAAGAGGCGAC 1669  
Db 1772 ACTCGCTTGTGATGAAGCGCTTCTGTTGGCCTGTGAGCGAAGAGTAAAGAGGCGAC 1831  
Qy 1670 TGGAGGTCTGCTGGGCTGCTCCGTTCCGCTGGGGGTGAAGAGAGCTTCTGCACTGGG 1729  
Db 1832 TGGAGGTCTGCTGGGCTGCTCCGTTCCGCTGGGGGTGAAGAGAGCTTCTGCACTGGG 1891  
Qy 1730 TCTCTCTGTGGGTGAGAGCTTAATGCAACACCCAGAGAGACCTCTGAGCGCTTCC 1789  
Db 1892 TCTCTCTGTGGGTGAGAGCTTAATGCAACACCCAGAGAGACCTCTGAGCGCTTCC 1951  
Qy 1790 ACTGTCTTGTGAGACTCAAGACAAAGATTGTTGCTTGGCATTTAAACAGCTTCCAG 1849  
Db 1952 ACTGTCTTGTGAGACTCAAGACAAAGATTGTTGCTTGGCATTTAAACAGCTTCCAG 2011  
Qy 1850 AAGTGTGCTTCCGATTAAACGAACCTTGAATACATCTTCTCTGCTCCAGC 1909  
Db 2012 AAGTGTGCTTCCGATTAAACGAACCTTGAATACATCTTCTCTGCTCCAGC 2071  
Qy 1910 ACTGTCCGATTGTCGGAATAATTCGGGTGATGTCAAGAGAGCTTCCCAAGAGTAGT 1969  
Db 2072 ACTGTCCGATTGTCGGAATAATTCGGGTGATGTCAAGAGAGCTTCCCAAGAGTAGT 2131  
Qy 1970 CCGCTGAGGCAATGTCTGTGTCTCTTATGATGCGGGATTAAGACCTCATTTAGAGAC 2029  
Db 2132 CCGCTGAGGCAATGTCTGTGTCTCTTATGATGCGGGATTAAGACCTCATTTAGAGAC 2191  
Qy 2030 AAGTGGAAATTTTGTCTGCAATGCTTGAACCCACACCTGCGGAGCTGAGACTGG 2089  
Db 2192 AAGTGGAAATTTTGTCTGCAATGCTTGAACCCACACCTGCGGAGCTGAGACTGG 2251  
Qy 2090 GCAGCAGCATCTGACAGAGCGGGCATGAAGACCTGTGTGCAAGCTGAGGCAATCCCA 2149  
Db 2252 GCAGCAGCATCTGACAGAGCGGGCATGAAGACCTGTGTGCAAGCTGAGGCAATCCCA 2311

Qy 2150 CTTGCAAGATACAGACCTGATGTTTGAATGCAAGATTAAACCTCTGTGTGACAGCAC 2209  
Db 2312 CTTGCAAGATACAGACCTGATGTTTGAATGCAAGATTAAACCTCTGTGTGACAGCAC 2371  
Qy 2210 TCTGAGAAATGCTATGAGCCCAACCGTAACTTAAGTCCCTCAACTTGGAGGCAACCC 2269  
Db 2372 TCTGAGAAATGCTATGAGCCCAACCGTAACTTAAGTCCCTCAACTTGGAGGCAACCC 2431  
Qy 2270 TGAAGAAAGAGATGAATGAGTGGCTGTGAAGCTTAAACACCCAAATGTTTGTGG 2329  
Db 2432 TGAAGAAAGAGATGAATGAGTGGCTGTGAAGCTTAAACACCCAAATGTTTGTGG 2491  
Qy 2330 AGCTTTAAGCTGATGCTGTGATGACCAATGCTGTAACTTGAAGATCTCCCAA 2389  
Db 2492 AGCTTTAAGCTGATGCTGTGATGACCAATGCTGTAACTTGAAGATCTCCCAA 2551  
Qy 2390 TCTTACGACCTCCCGCAGCTGAATCTGAGCTGAGAGAAACAAAGTGAAGAC 2449  
Db 2552 TCTTACGACCTCCCGCAGCTGAATCTGAGCTGAGAGAAACAAAGTGAAGAC 2611  
Qy 2450 AGGAGTAATGCTCTCAGTATGCTTGAAGTCTCCAGTGCCTGCAAGAGCTGA 2509  
Db 2612 AGGAGTAATGCTCTCAGTATGCTTGAAGTCTCCAGTGCCTGCAAGAGCTGA 2671  
Qy 2510 TACTGAGAGACTGTGGATCAACAGCCAGGTTGCAAGTCTGGCTCAGCCCTGCTCA 2569  
Db 2672 TACTGAGAGACTGTGGATCAACAGCCAGGTTGCAAGTCTGGCTCAGCCCTGCTCA 2731  
Qy 2570 GCAACCGGAGCTGACACACCTGTGCTATCCAAACAGCCTGGGGAACGAAGGTGA 2629  
Db 2732 GCAACCGGAGCTGACACACCTGTGCTATCCAAACAGCCTGGGGAACGAAGGTGA 2791  
Qy 2630 ATCTACTGTGTGATCATGAGCTTCCCACTGATCTGCAAGAGCTGATGCTGAATC 2689  
Db 2792 ATCTACTGTGTGATCATGAGCTTCCCACTGATCTGCAAGAGCTGATGCTGAATC 2851  
Qy 2690 AAGTCCACTCTGAGACAGGCTGCTGTGTTTCTTGCACCTGCTTATGGGTACTCAT 2749  
Db 2852 AAGTCCACTCTGAGACAGGCTGCTGTGTTTCTTGCACCTGCTTATGGGTACTCAT 2911  
Qy 2750 GGTGACGACCTGAGCCTTATGATGAACCTGTGGAAGACAAATGGCGTGAAGCTTCTGT 2809  
Db 2912 GGTGACGACCTGAGCCTTATGATGAACCTGTGGAAGACAAATGGCGTGAAGCTTCTGT 2971  
Qy 2810 GCGAGTCAATGAGAAACATCTTGTCACTTCCAGAGCTGGAAGTTGGTAAAGTCAATC 2869  
Db 2972 GCGAGTCAATGAGAAACATCTTGTCACTTCCAGAGCTGGAAGTTGGTAAAGTCAATC 3031  
Qy 2870 TCAACCGCGCTGTGTGAGAGTCTGTCTGTGTGATCTGAGAGAGACACTGAAGA 2929  
Db 3032 TCAACCGCGCTGTGTGAGAGTCTGTCTGTGTGATCTGAGAGAGACACTGAAGA 3091  
Qy 2930 GCTGTGATCTACGAGCAATGCGCTGGGTGAACGAGTGGGTGCTGCGTGGCGAGGAC 2989  
Db 3092 GCTGTGATCTACGAGCAATGCGCTGGGTGAACGAGTGGGTGCTGCGTGGCGAGGAC 3151  
Qy 2990 TGAAGCAAAAGAACAGTGTCTGACGAGACTCGGTTGAAGGCAATGTGACTGACTTCTG 3049  
Db 3152 TGAAGCAAAAGAACAGTGTCTGACGAGACTCGGTTGAAGGCAATGTGACTGACTTCTG 3211  
Qy 3050 ATTGCTGTGAGCACTTCTCTTGGCCCTTCTCTGCAACCGGATGTGACAGTCTAAAC 3109  
Db 3212 ATTGCTGTGAGCACTTCTCTTGGCCCTTCTCTGCAACCGGATGTGACAGTCTAAAC 3271  
Qy 3110 TGGTGCAGAAATTAACCTTCAAGTCCAAAGGAATGATGAAGCTGTTGCGCTTGGCTGTC 3169  
Db 3272 TGGTGCAGAAATTAACCTTCAAGTCCAAAGGAATGATGAAGCTGTTGCGCTTGGCTGTC 3331  
Qy 3170 CCAAGCTTAACCTTACAGATTAATTTGGCTGTGGAATGCAAGTACCTCTGTGCAAAATGA 3229  
Db 3332 CCAAGCTTAACCTTACAGATTAATTTGGCTGTGGAATGCAAGTACCTCTGTGCAAAATGA 3391



QY 1425 GAACATCTCTTCCAGACAGCACTGTAGAGATCTACACCTTCTTCCACTCAGTCT 1484  
 DB 1530 GAAACATCTCTTCCAGACAGCACTGTAGAGATCTACACCTTCTTCCACTCAGTCT 1589  
 QY 1485 CAGAGATCTTGTGCGCCTTGTACTACGTGTAAAGGCGCTGGAATGAGCCAGCTCT 1544  
 DB 1590 CAGAGATCTTGTGCGCCTTGTACTACGTGTAAAGGCGCTGGAATGAGCCAGCTCT 1649  
 QY 1545 CCGCCCTGTAGAGATGAGAGCAAGAGGCTCATGAGGCTTAAACAGGCGCTTCCA 1604  
 DB 1650 CCGCCCTGTAGAGATGAGAGCAAGAGGCTCATGAGGCTTAAACAGGCGCTTCCA 1709  
 QY 1605 TATCCACTGCTTGTGATGAAAGCTTCTTGTGTGCGCTGTAGAGGAGAGATGAGAG 1664  
 DB 1710 TATCCACTGCTTGTGATGAAAGCTTCTTGTGTGCGCTGTAGAGGAGAGATGAGAG 1769  
 QY 1665 GCGACTGAGAGTCTGTGCGGCTGTCCGCTTCCCTGCGGAGTGAAGCAGAGCTTGTGCA 1724  
 DB 1770 GCGACTGAGAGTCTGTGCGGCTGTCCGCTTCCCTGCGGAGTGAAGCAGAGCTTGTGCA 1829  
 QY 1725 CTGGGCTCTCTGTGTGGGTGAGAGCTTAAATGACCAACCCAGAGAGACCTTGAAGCG 1784  
 DB 1830 CTGGGCTCTCTGTGTGGGTGAGAGCTTAAATGACCAACCCAGAGAGACCTTGAAGCG 1889  
 QY 1785 CTTCACATGCTTCTTGTGAGACTCAAGACAAAGAGTTGTGCTTGTGAGATTAACAGCTT 1844  
 DB 1890 CTTCACATGCTTCTTGTGAGACTCAAGACAAAGAGTTGTGCTTGTGAGATTAACAGCTT 1949  
 QY 1845 CCAAGAGTGTGGCTTCCGATTAACCAAGACCTTGAATGATGATCTTCTTGTGCTT 1904  
 DB 1950 CCAAGAGTGTGGCTTCCGATTAACCAAGACCTTGAATGATGATCTTCTTGTGCTT 2009  
 QY 1905 CCAAGATCTGTCCGTATTTGCGGAAATTTGGGTGATGTCAAAAGGATTTTCCCAAGAG 1964  
 DB 2010 CCAAGATCTGTCCGTATTTGCGGAAATTTGGGTGATGTCAAAAGGATTTTCCCAAGAG 2069  
 QY 1965 TGAAGTCGCTGAGAGATGTCCTGTGGTCCCTCATAGAGATGCGGAGTAAAGCCCTGATGGA 2024  
 DB 2070 TGAAGTCGCTGAGAGATGTCCTGTGGTCCCTCATAGAGATGCGGAGTAAAGCCCTGATGGA 2129  
 QY 2025 GAGAGATGTGGAGAAATTTCTGTCTCAATGCTTGGCAACCAACCACTGCGGAGCTGGA 2084  
 DB 2130 GAGAGATGTGGAGAAATTTCTGTCTCAATGCTTGGCAACCAACCACTGCGGAGCTGGA 2189  
 QY 2085 CTTGGGAGAGAGATCTCTGACAGAGCGGCGCATGAAGACCTGTGTGCCAAGCTGAGGCA 2144  
 DB 2190 CTTGGGAGAGAGATCTCTGACAGAGCGGCGCATGAAGACCTGTGTGCCAAGCTGAGGCA 2249  
 QY 2145 TCCCACTGTGAGATGACAGACCTGATGTTTGAATGACAGATTAACCCCTGTGTGCA 2204  
 DB 2250 TCCCACTGTGAGATGACAGACCTGATGTTTGAATGACAGATTAACCCCTGTGTGCA 2309  
 QY 2205 GCACTCTGTGAGAAATGTCATGAGCAACCGTAACCTTAAGATCCCTCAATTGGAGGAGC 2264  
 DB 2310 GCACTCTGTGAGAAATGTCATGAGCAACCGTAACCTTAAGATCCCTCAATTGGAGGAGC 2369  
 QY 2265 CCACTTGAAGAGAGATGTAAGATGCGGTGTAAGCTTAAACACCCAAATGTTT 2324  
 DB 2370 CCACTTGAAGAGAGATGTAAGATGCGGTGTAAGCTTAAACACCCAAATGTTT 2429  
 QY 2325 GTTGAAGCTTGTGAGGCTGAGATTTGCTGTGAGTTGACCAATGCTGTAAAGATGTC 2384  
 DB 2430 GTTGAAGCTTGTGAGGCTGAGATTTGCTGTGAGTTGACCAATGCTGTAAAGATGTC 2489  
 QY 2385 CCAAAATCTTACAGACTCCCGAGCTGAAATCTCTGAGCTGTGAGAGAAACAAGTGAAC 2444  
 DB 2490 CCAAAATCTTACAGACTCCCGAGCTGAAATCTCTGAGCTGTGAGAGAAACAAGTGAAC 2549  
 QY 2445 AGACCAAGAGAGATGATGCTCTCAAGTATGCTTGAAGAGTCTCCCAAGTGGCGCTTCAAGAA 2504  
 DB 2550 AGACCAAGAGAGATGATGCTCTCAAGTATGCTTGAAGAGTCTCCCAAGTGGCGCTTCAAGAA 2609  
 QY 2505 GGTGATATCTGAGAGACTGTGATCAACAGCAACGGGTTGCCAGAGTCTGCGCTCAAGCCT 2564

DB 2610 GGTGATATCTGAGAGACTGTGGATCAAGCAACGGGTTGCCAGAGTCTGGCTCAGCCCT 2669  
 QY 2565 GGTGAGCAACCGGAGCTTGAACACACTGTGCTTATCAACAACAGCTTGGGGAACGAAGG 2624  
 DB 2670 GGTGAGCAACCGGAGCTTGAACACACTGTGCTTATCAACAACAGCTTGGGGAACGAAGG 2729  
 QY 2625 TGTAAATATCTGTGTGATCATGAGGCTTCCCACTGTGTGTGAGAGGCTGATGCT 2684  
 DB 2730 TGTAAATATCTGTGTGATCATGAGGCTTCCCACTGTGTGTGAGAGGCTGATGCT 2789  
 QY 2685 GAATAGTGCACCTGAGACAGGCTGTGAGTGTGTTTCTTCAACTTGGCTTAATGGTAA 2744  
 DB 2790 GAATAGTGCACCTGAGACAGGCTGTGAGTGTGTTTCTTCAACTTGGCTTAATGGTAA 2849  
 QY 2745 CTGATGCTGAGAGCACTGAGCTTGAAGTAAACCTTGTGAGAAACAATGCGTGAAGCT 2804  
 DB 2850 CTGATGCTGAGAGCACTGAGCTTGAAGTAAACCTTGTGAGAAACAATGCGTGAAGCT 2909  
 QY 2805 TCTGTGAGAGGCTCATGAGAGAAACCAATCTGTATCTCCAGAGACTGAGTGTGTAAGTG 2864  
 DB 2910 TCTGTGAGAGGCTCATGAGAGAAACCAATCTGTATCTCCAGAGACTGAGTGTGTAAGTG 2969  
 QY 2865 TCATCTCAACCGCGGCTGTGAGAGTCTGCTGTGTGATCTTGAAGAGAGACACACT 2924  
 DB 2970 TCATCTCAACCGCGGCTGTGAGAGTCTGCTGTGTGATCTTGAAGAGAGACACACT 3029  
 QY 2925 GAAAGCTGTGATCTCAAGCAACATGCTGTGAGAGAGGAGTGTGCTGCGCTGTGCA 2984  
 DB 3030 GAAAGCTGTGATCTCAAGCAACATGCTGTGAGAGAGGAGTGTGCTGCGCTGTGCA 3089  
 QY 2985 GGAATGAAAGCAAAAGAACATGTTCTGACAGAGCTGCGGTTGAAGGATGTGACTGAC 3044  
 DB 3090 GGAATGAAAGCAAAAGAACATGTTCTGACAGAGCTGCGGTTGAAGGATGTGACTGAC 3149  
 QY 3045 TTCTGATGCTGTGAGGAGCACTCTCTTGGCCCTTCTGCAACCGGATCTGACAGTCT 3104  
 DB 3150 TTCTGATGCTGTGAGGAGCACTCTCTTGGCCCTTCTGCAACCGGATCTGACAGTCT 3209  
 QY 3105 AAACCTGTGAGAAATTAATCTTCAATGCTCCCAAGAAATGATGAAGCTGTGCGCTTGTG 3164  
 DB 3210 AAACCTGTGAGAAATTAATCTTCAATGCTCCCAAGAAATGATGAAGCTGTGCGCTTGTG 3269  
 QY 3165 CTGTCCCACTGTCTAATTAAGATTAATGAGTGTGAGAAATGAGCAATCTGTGCAAT 3224  
 DB 3270 CTGTCCCACTGTCTAATTAAGATTAATGAGTGTGAGAAATGAGCAATCTGTGCAAT 3329  
 QY 3225 AAGGAAGCTGTGAGAGAGAGTCAAGCTCAAGCCCGAGTCTGTAATTGAACGATGTTG 3284  
 DB 3330 AAGGAAGCTGTGAGAGAGAGTCAAGCTCAAGCCCGAGTCTGTAATTGAACGATGTTG 3389  
 QY 3285 GCATTCCTTGTGATGAAGTGAACCG 3308  
 DB 3390 GCATTCCTTGTGATGAAGTGAACCG 3413

RESULT 6  
 AAL47135  
 ID AAL47135 standard; DNA; 5859 BP.  
 XX AAL47135;  
 DT 20-AUG-2002 (first entry)  
 XX  
 DE Pyrin domain containing protein NALP6/Py12 coding sequence.  
 XX  
 KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;  
 KW antiarteriosclerotic; antiproliferative; antibacterial; virucide;  
 KW neuroprotective; antiarthritic; antirheumatic; antiaesthetic;  
 KW nephrotropic; osteoplastic; neotrophic; intercellular signal transduction;  
 KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;  
 KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;  
 KW osteoarthritis; glomerulonephritis; gene; db.

XX Unidentified.  
OS  
XX WO200240668-A2.  
XX  
XX 23-MAY-2002.  
XX  
XX 30-OCT-2001; 2001WO-BP012545.  
XX  
XX 15-NOV-2000; 2000DE-01056587.  
XX 30-NOV-2000; 2000DE-01059595.  
XX  
XX (ABOT-) APOTEC RES & DEV LTD.  
XX  
XX Tschoopp J, Martinon F;  
XX  
XX WPI; 2002-427093/45.  
XX P-PSDB; AA017863.  
XX  
XX New DNA encoding protein with pyrin domain, useful for treating diseases  
XX involving impaired signal transduction, particularly inflammation, also  
XX proteins and antibodies.  
XX  
XX Claim 5; Fig 1; 116pp; German.  
XX  
XX The present invention relates the DNA and their encoded proteins, where  
XX the proteins contain at least one PYD (pyrin) domain. These can be used  
XX to treat diseases associated with impaired intracellular signal  
XX transduction, particularly inflammation such as psoriasis,  
XX arteriosclerosis, bacterial or viral infections (particularly meningitis  
XX and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,  
XX sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's  
XX and Parkinson's diseases. The present sequence is a coding sequence of  
XX the invention  
XX  
XX Sequence 5859 BP; 1512 A; 1443 C; 1530 G; 1374 T; 0 U; 0 Other;  
XX  
XX Query Match 69.5%; Score 2803; DB 6; Length 5859;  
XX Best Local Similarity 99.9%; Pred. No. 0;  
XX Matches 2903; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
XX  
XX 290 AAATTTCACAACTATGGAACAAGAGTGCACAGCAGCAGACAGACAGAAACAAGAA 349  
XX 2387 AAATTTCACAACTATGGAACAAGAGTGCACAGCAGCAGACAGACAGAAACAAGAA 2446  
XX 350 TTTTCAAGCTATGGAACAAGAGTGCACAGCAGCAGACAGACAGAAACAAGAACTG 409  
XX 2447 TTTTCAAGCTATGGAACAAGAGTGCACAGCAGCAGACAGACAGAAACAAGAACTG 2506  
XX 410 GAGGTGACATGAGGACTACAAAGATCAGTGATGACCAAAATTCGCTGAGAGAGAGATG 469  
XX 2507 GAGGTGACATGAGGACTACAAAGATCAGTGATGACCAAAATTCGCTGAGAGAGAGATG 2566  
XX 470 TACGTGATGTTTGAAGAACTGCTGCTGACTGCGCGGAAATGCAAACTGCTGCTG 529  
XX 2567 TACGTGATGTTTGAAGAACTGCTGCTGACTGCGCGGAAATGCAAACTGCTGCTG 2626  
XX 530 CTTTGTGATTCAGACCGGTGGGCTTCGCGCTCGACCGGTGTTTGAACGGAAGTCA 589  
XX 2627 CTTTGTGATTCAGACCGGTGGGCTTCGCGCTCGACCGGTGTTTGAACGGAAGTCA 2686  
XX 590 GAATTGGGAATTCGCTTACGACAGAGATCGTCTGCTGCGGCGCAAGTGGACTCT 649  
XX 2687 GAATTGGGAATTCGCTTACGACAGAGATCGTCTGCTGCGGCGCAAGTGGACTCT 2746  
XX 650 ACCAGGAATGTTCTCTACGCTTCTTCTCTCCCGTTAGAGAGATGCAAGCGGAAGAG 709  
XX 2747 ACCAGGAATGTTCTCTACGCTTCTTCTCTCCCGTTAGAGAGATGCAAGCGGAAGAG 2806  
XX 710 AGAGCAGTGTCAACAGATTCATCTCCAGGAGTGGCCGAGCTCCAGGCTCCGGTGAAG 769  
XX 2807 AGAGCAGTGTCAACAGATTCATCTCCAGGAGTGGCCGAGCTCCAGGCTCCGGTGAAG 2866  
XX  
XX

QY 770 AGATCATGTCCCGACAGAAAGCTGTGTTCATCATATGACGTTTGCATGACCTGGGCT 829  
DB 2867 AGATCATGTCCCGACAGAAAGCTGTGTTCATCATATGACGTTTGCATGACCTGGGCT 2926  
QY 830 CTGTCTCAACATGACCAAAAGCTCTGCAAAAGCTGGGCTGAGAGACGCTCCGTTCA 889  
DB 2927 CTGTCTCAACATGACCAAAAGCTCTGCAAAAGCTGGGCTGAGAGACGCTCCGTTCA 2986  
QY 890 CCTCATACGACGTCTGCTGAGGAAGGTCGCTCCGAGTCTCTGATCGTCAACG 949  
DB 2987 CCTCATACGACGTCTGCTGAGGAAGGTCGCTCCGAGTCTCTGATCGTCAACG 3046  
QY 950 TCAGAGAGTGGGACAGAGAGCTCAAGTCAGAGTGTGTCTCCCGTTACCTGTTAG 1009  
DB 3047 TCAGAGAGTGGGACAGAGAGCTCAAGTCAGAGTGTGTCTCCCGTTACCTGTTAG 3106  
QY 1010 TTAGAGGAATTCGCGGGAACAAGAAATCCATTGCTCTTGAAGCGGGATTGGTAGC 1069  
DB 3107 TTAGAGGAATTCGCGGGAACAAGAAATCCATTGCTCTTGAAGCGGGATTGGTAGC 3166  
QY 1070 ATCAGAGACACAAAGGTTGGTGGATCATGAACAAACCGTACGCTGCGACAGTGGC 1129  
DB 3167 ATCAGAGACACAAAGGTTGGTGGATCATGAACAAACCGTACGCTGCGACAGTGGC 3226  
QY 1130 AGTGCCCGCGCGTGGGCTCTCTCATCTGCGTGGGCTGACGCTGACAGACGTTGGGCG 1189  
DB 3227 AGTGCCCGCGCGTGGGCTCTCTCATCTGCGTGGGCTGACGCTGACAGACGTTGGGCG 3286  
QY 1190 AGAGGTGGCCCGCTTCAACCAACGCTCACAGGCTGACAGCGCTTTGTGTTCATC 1249  
DB 3287 AGAGGTGGCCCGCTTCAACCAACGCTCACAGGCTGACAGCGCTTTGTGTTCATC 3346  
QY 1250 AGCTCACCCCTGAGGGGTGGTCCGGGCTGTCTCAATCTGAGAGAAAGATTGCTCGA 1309  
DB 3347 AGCTCACCCCTGAGGGGTGGTCCGGGCTGTCTCAATCTGAGAGAAAGATTGCTCGA 3406  
QY 1310 AGCGCTTCTGCGCTGATGAGTGTGAGAGAGAGTGGAAATAGGAATCAGTGTATGATG 1369  
DB 3407 AGCGCTTCTGCGCTGATGAGTGTGAGAGAGAGTGGAAATAGGAATCAGTGTATGATG 3466  
QY 1370 AGCACTCATGTGTTCAAGAGCTCGGGAGTGTGAGCTCCGTCTGTGTTCATATGAACA 1429  
DB 3467 AGCACTCATGTGTTCAAGAGCTCGGGAGTGTGAGCTCCGTCTGTGTTCATATGAACA 3526  
QY 1430 TCTCTTCCGACAGAGCACTGTGAGAGTACTACACTTCTTCCACTGCTCCAGG 1489  
DB 3527 TCTCTTCCGACAGAGCACTGTGAGAGTACTACACTTCTTCCACTGCTCCAGG 3586  
QY 1490 ACTTCTGTCCGCTTGTACTACGTGTGAGAGGCTGGAATGAGCCAGCTCTGCC 1549  
DB 3587 ACTTCTGTCCGCTTGTACTACGTGTGAGAGGCTGGAATGAGCCAGCTCTGCC 3646  
QY 1550 CTCTGTACGTTGAGAGCAAAAGGTTCATGAGCTTTAAACAGGCAAGCTTCAATATCC 1609  
DB 3647 CTCTGTACGTTGAGAGCAAAAGGTTCATGAGCTTTAAACAGGCAAGCTTCAATATCC 3706  
QY 1610 ACTGCTTTGAGATGAAGGTTCTTGTGTCCTGAGGCTCGTGAAGAGAGTAAAGAGGAC 1669  
DB 3707 ACTGCTTTGAGATGAAGGTTCTTGTGTCCTGAGGCTCGTGAAGAGAGTAAAGAGGAC 3766  
QY 1670 TGAAGTCTGCTGGGCTGTCCGTTCCCTGCGGGGTGAAGCAGAACTTGTGACTGGG 1729  
DB 3767 TGAAGTCTGCTGGGCTGTCCGTTCCCTGCGGGGTGAAGCAGAACTTGTGACTGGG 3826  
QY 1730 TCTCTGTGTGGTCAAGAGCTTAATGCAACCCCAAGAGACACCTTGACGCTTCC 1789  
DB 3827 TCTCTGTGTGGTCAAGAGCTTAATGCAACCCCAAGAGACACCTTGACGCTTCC 3886  
QY 1790 ACTGCTTTTGAAGACTCAAGACAAAGTGTGTGCTTGGCTTAAACAGCTTCAAG 1849  
DB 3887 ACTGCTTTTGAAGACTCAAGACAAAGTGTGTGCTTGGCTTAAACAGCTTCAAG 3946  
QY 1850 AAGTGTGGCTTCCGATTAAACCAAGACTGGAATTGATGATCTTCTGCTCCAGC 1909

```

Db 3947 AAGTGGCTTCGATTACGAACTGGAATTGATGACATCTTCTTCTGCTCCAGC 4006
Qy 1910 ACTGTCCTATTGGGAAAAATTCGGGTGATGTCAAAGGATCTTCCCAAGATGAGT 1969
Db 4007 ACTGTCCTATTGGGAAAAATTCGGGTGATGTCAAAGGATCTTCCCAAGATGAGT 4066
Qy 1970 CCGGTGAGGCAATGTCCTGTCCTCTATGGAATCGGGATGAAGCCCTCATTTGAGGAC 2029
Db 4067 CCGGTGAGGCAATGTCCTGTCCTCTATGGAATCGGGATGAAGCCCTCATTTGAGGAC 4126
Qy 2030 AGTGGGAAGATTTCTGCTCCATGCTTGGAACCCACCACTGGGAGCTGACCTGG 2089
Db 4127 AGTGGGAAGATTTCTGCTCCATGCTTGGAACCCACCACTGGGAGCTGACCTGG 4186
Qy 2090 GCAGCAGCATCTGTCAGAGCGGSCCATGAGAACCTTGTCGCAAGCTGAGCATCCCA 2149
Db 4187 GCAGCAGCATCTGTCAGAGCGGSCCATGAGAACCTTGTCGCAAGCTGAGCATCCCA 4246
Qy 2150 CCTGGAATATCAGACCCCTGATGTTTAAATGCAAGATTAACCTGTCGAGCACC 2209
Db 4247 CCTGGAATATCAGACCCCTGATGTTTAAATGCAAGATTAACCTGTCGAGCACC 4306
Qy 2210 TCTGGAATATCCTCATGAGCCGTAACCTTAAGATCCCTCACTTGGAGGACCCACC 2269
Db 4307 TCTGGAATATCCTCATGAGCCGTAACCTTAAGATCCCTCACTTGGAGGACCCACC 4366
Qy 2270 TGAAGAAAGAGATGTAAGATGCGGTGTAAGCCTTAAACACCCAAATGTTTGG 2329
Db 4367 TGAAGAAAGAGATGTAAGATGCGGTGTAAGCCTTAAACACCCAAATGTTTGG 4426
Qy 2330 AGCTTTTGAAGGCTGATGCTGTCGATGTAACCATGCTGTTAACCCTGAAGATCTCCAA 2389
Db 4427 AGCTTTTGAAGGCTGATGCTGTCGATGTAACCATGCTGTTAACCCTGAAGATCTCCAA 4486
Qy 2390 TCCCTAAGACCTCCCCAGCCTGAATCTCTGAGCCTGGCAGAGAAACAAGTGAAGACC 2449
Db 4487 TCCCTAAGACCTCCCCAGCCTGAATCTCTGAGCCTGGCAGAGAAACAAGTGAAGACC 4546
Qy 2450 AAGGAGTAATGCTCTCAGTGAATGCTTGAAGATCTCCAGTGCAGCCCTGCAAGAGCTGA 2509
Db 4547 AAGGAGTAATGCTCTCAGTGAATGCTTGAAGATCTCCAGTGCAGCCCTGCAAGAGCTGA 4606
Qy 2510 TACTGAGAGATCTGTGGCATCAAGCCAGGGTTGCGAAGATTTGGCCTTCTGCTCA 2569
Db 4607 TACTGAGAGATCTGTGGCATCAAGCCAGGGTTGCGAAGATTTGGCCTTCTGCTCA 4666
Qy 2570 GCAACCGGAGCTTGACACACCTGTGCTATCCAAACAAGCCTGGGGAACGAAGGTGTA 2629
Db 4667 GCAACCGGAGCTTGACACACCTGTGCTATCCAAACAAGCCTGGGGAACGAAGGTGTA 4726
Qy 2630 ATCTACTGTGTCATCATGAGGCTTCCCACTGATGCTGCAAGAGCTGATGCTGAATC 2689
Db 4727 ATCTACTGTGTCATCATGAGGCTTCCCACTGATGCTGCAAGAGCTGATGCTGAATC 4786
Qy 2690 AGTGCACCTTGACACAGGCTGCTGTTTCTTGCACTTGCGCTTATGAGTAACTCAT 2749
Db 4787 AGTGCACCTTGACACAGGCTGCTGTTTCTTGCACTTGCGCTTATGAGTAACTCAT 4846
Qy 2750 GGCCTGAGCACTGAGCCTTATGATGAACCCCTGTGGAAGCAATGCGCTGAAGCTTCTGT 2809
Db 4847 GGCCTGAGCACTGAGCCTTATGATGAACCCCTGTGGAAGCAATGCGCTGAAGCTTCTGT 4906
Qy 2810 GCGAGGTATGAGAGAACCATCTTGTCACTTCAGAGACCTGAGATTGGTAAAGTGTCAATC 2869
Db 4907 GCGAGGTATGAGAGAACCATCTTGTCACTTCAGAGACCTGAGATTGGTAAAGTGTCAATC 4966
Qy 2870 TCACCGCGCGCTGCTGTGAGATCTGTCTGTGTGATCTCGAGAGCAGACACTGAAGA 2929
Db 4967 TCACCGCGCGCTGCTGTGAGATCTGTCTGTGTGATCTCGAGAGCAGACACTGAAGA 5026
Qy 2930 GCGTGGATCTCAGGAGCAATGCGCTGGGTGACGGGTGGGTTGCTGCGCTGCGAGGAGAC 2989

```

```

Db 5027 GCTGTGATCTCAGGAGCAATGCTCCGTGGGTGACGGTGGGGTGTGTCGACTGTGAGGAGAC 5086
Qy 2990 TGAAGCAAAAGAAAGATGTTTTCGACGAGACTCGGGTTGAAGGCATGTGTGACTTGTG 3049
Db 5087 TGAAGCAAAAGAAAGATGTTTTCGACGAGACTCGGGTTGAAGGCATGTGTGACTTGTG 5146
Qy 3050 ATTGCTGTGAGGCACTCTCCCTTGGCCCTTCTGCAACCGGCACTGACAGATCTAAAC 3109
Db 5147 ATTGCTGTGAGGCACTCTCCCTTGGCCCTTCTGCAACCGGCACTGACAGATCTAAAC 5206
Qy 3110 TGTGTGCAATTAATCTTCAAGTCCCAAGAGATGATGAGCTGTGTGGCCTTGTGCTGTC 3169
Db 5207 TGTGTGCAATTAATCTTCAAGTCCCAAGAGATGATGAGCTGTGTGGCCTTGTGCTGTC 5266
Qy 3170 CCAGGTCTTAATCTTCAAGATTTGG 3194
Db 5267 CCAGGTCTTAATCTTCAAGATTTGG 5291

```

```

RESULT 7
AAL47131
ID AAL47131 standard; DNA; 6939 BP.
XX
XX AAL47131;
XX
XX 20-AUG-2002 (first entry)
XX
XX Pyrin domain containing protein NALP5/Py8-hs coding sequence.
DE
XX Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
XX antiarteriosclerotic; antipsoriatic; antibacterial; virocid;
XX neuroprotective; antiarthritic; antirheumatic; antiaesthetic;
XX nephroprotective; osteoprotic; nootropic; intracellular signal transduction;
XX inflammation; Alzheimer's disease; infection; psoriasis; asthma;
XX arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
XX osteoarthritis; glomerulonephritis; gene; ds.
XX
XX Unidentified.
OS
XX
XX W0200240668-A2.
XX
XX 23-MAY-2002.
XX
XX 30-OCT-2001; 2001WO-EP012545.
XX
XX 15-NOV-2000; 2000DE-01056687.
XX
XX 30-NOV-2000; 2000DE-01059595.
XX
XX (APOF-) APOTECH RES & DEV LTD.
XX
XX Tschoopp J, Martinon F;
XX
XX MPI; 2002-427093/45.
XX
XX P-P8DB; AAO17859.
XX
XX New DNA encoding protein with pyrin domain, useful for treating diseases
XX PT involving impaired signal transduction, particularly inflammation, also
XX PT proteins and antibodies.
XX
XX Claim 5; Fig 1; 116pp; German.
XX
XX The present invention relates the DNA and their encoded proteins, where
XX CC the proteins contain at least one PYD (pyrin) domain. These can be used
XX CC to treat diseases associated with impaired intracellular signal
XX CC transduction, particularly inflammation such as psoriasis,
XX CC arteriosclerosis, bacterial or viral infections (particularly meningitis
XX CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
XX CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
XX CC and Parkinson's diseases. The present sequence is a coding sequence of
XX CC the invention.
XX
XX Sequence 6939 BP; 1791 A; 1719 C; 1801 G; 1628 T; 0 U; 0 Other;
XX
XX

```



Query Match	69.5%	Score 2803	DB 6	Length 6939
Beet Local Similarity	99.9%	Prod. No. 0	Mismatches 2	Indels 0
Matches 2903	Conservative			Gaps 0
QY	290	AAATTTCAACAGCTATGGACAAGAAGGTGGCCACAGCAGCAGACAGACAAGAACAAGAA	349	
Db	3467	AAATTTCAACAGCTATGGACAAGAAGGTGGCCACAGCAGCAGACAGACAAGAACAAGAA	3528	
QY	350	TTTTCACAAGCTATGGACAAGAAGGTGGCCACAGCAGCAGACAGACAAGAACAAGCA	409	
Db	3527	TTTTCACAAGCTATGGACAAGAAGGTGGCCACAGCAGCAGACAGACAAGAACAAGCA	3588	
QY	410	GAGGTGACACATGGGACTTACAAAGATCACTGATATACCAAAATTCGCTGAGAGAGATG	469	
Db	3587	GAGGTGACACATGGGACTTACAAAGATCACTGATATACCAAAATTCGCTGAGAGAGATG	3648	
QY	470	TACGTCGATGTTTGAACAACTGCTGCTGACTGGCCGGAATATGMAACGTTGGCTGGT	529	
Db	3647	TACGTCGATGTTTGAACAACTGCTGCTGACTGGCCGGAATATGMAACGTTGGCTGGT	3706	
QY	530	CTTTTGATTCAGACCGGTGGGCTTCCGGCTCCGACCGGTGGATTCGACCGAAAGTCAG	589	
Db	3707	CTTTTGATTCAGACCGGTGGGCTTCCGGCTCCGACCGGTGGATTCGACCGAAAGTCAG	3768	
QY	590	GAATTTGGGAATTCGGCTCTAGCCAGAAAGATCTGCTGCTGGCCGCAAGTGGACTCT	649	
Db	3767	GAATTTGGGAATTCGGCTCTAGCCAGAAAGATCTGCTGCTGGCCGCAAGTGGACTCT	3828	
QY	650	ACCAAGGAATGTTCTCTACGCTCTTCTCCCGCTTAGAGAGATGCAAGCGAAGAG	709	
Db	3827	ACCAAGGAATGTTCTCTACGCTCTTCTCCCGCTTAGAGAGATGCAAGCGAAGAG	3888	
QY	710	AGAGCAGTGTCAACAGATTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGCTGACGG	769	
Db	3887	AGAGCAGTGTCAACAGATTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGCTGACGG	3948	
QY	770	AGATCATGTCCCGACCAAAAGGCTGTGTTTCATATTGACGTTTCATGACTCGGCT	829	
Db	3947	AGATCATGTCCCGACCAAAAGGCTGTGTTTCATATTGACGTTTCATGACTCGGCT	4008	
QY	830	CTGTCCTCAACAATGACCAAAAGCTCTGCAAAAGCTGGGCTGAGAAAGAGCTCCGTTCA	889	
Db	4007	CTGTCCTCAACAATGACCAAAAGCTCTGCAAAAGCTGGGCTGAGAAAGAGCTCCGTTCA	4068	
QY	890	CCCTCATACGAGTCTGCTGAGGAAGGCTCTGCTCCCTGAGTCTTCTCTGATTCGTCA	949	
Db	4067	CCCTCATACGAGTCTGCTGAGGAAGGCTCTGCTCCCTGAGTCTTCTCTGATTCGTCA	4128	
QY	950	TCAGAGACGTGGGCAACAGAAAGCTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTTAG	1009	
Db	4127	TCAGAGACGTGGGCAACAGAAAGCTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTTAG	4188	
QY	1010	TTAGAGGAATCTCCGGGGACAAGAAATCCACTTGCTCTTGAGGCGGGAATGGTGA	1065	
Db	4187	TTAGAGGAATCTCCGGGGACAAGAAATCCACTTGCTCTTGAGGCGGGAATGGTGA	4248	
QY	1070	ATCAGAAACACAAGGTTGGCTGGATCATGAACMACCGTGAAGTGTCTGCACCACTG	1129	
Db	4247	ATCAGAAACACAAGGTTGGCTGGATCATGAACMACCGTGAAGTGTGTCTGCACCACTG	4306	
QY	1130	AGTGTCCCGCGTGGGCTTCTCATCTGCGTGGCCCTGACGTGACGACGCTTGTGTTT	1189	
Db	4307	AGTGTCCCGCGTGGGCTTCTCATCTGCGTGGCCCTGACGTGACGACGCTTGTGTTT	4366	
QY	1190	AGAGGTGCGCCCTTTCAACCAAGATCAAGGCTGACAGCGCTTGTGTTTTCATC	1244	
Db	4367	AGAGGTGCGCCCTTTCAACCAAGATCAAGGCTGACAGCGCTTGTGTTTTCATC	4428	
QY	1250	AGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTGGAGAAAGATGTTCTG	1309	
Db	4427	AGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTGGAGAAAGATGTTCTG	4488	
QY	1310	AGCGCTTCTGCCGATGAGCTGTGAGGAGATGTGAAATAGAAAGTCAAGTGTGATG	1365	

[illegible]



```

Db 5567 TCCTTACGACCTCCCGCCAGCTTAAATCTCTGAGCCTGCGAGAGAAACAAAGTGACAGACC 5626
QY 2450 AGGAGTAATGCTCTCAGTGAATGCTTGAAGTCCCGAGTGGCGCCCTGAGAAAGCTGA 2509
Db 5627 AGGAGTAATGCTCTCAGTGAATGCTTGAAGTCCCGAGTGGCGCCCTGAGAAAGCTGA 5686
QY 2510 TACTGAGAGCTGTGGCATCAGAGCCAGGGTGGCCAGATGTCGAGCTTGCCTGTCA 2569
Db 5687 TACTGAGAGCTGTGGCATCAGAGCCAGGGTGGCCAGATGTCGAGCTTGCCTGTCA 5746
QY 2570 GCAACCGAGCTTGAACACCTGTGCTATCCAAACAAGCTTGGGAAACGAAGGTGTA 2629
Db 5747 GCAACCGAGCTTGAACACCTGTGCTATCCAAACAAGCTTGGGAAACGAAGGTGTA 5806
QY 2630 ATCTACTGTGTGATCCAGTGAAGGCTTCCCGATGATGTCGAGAGGCTGATGCTGAATC 2689
Db 5807 ATCTACTGTGTGATCCAGTGAAGGCTTCCCGATGATGTCGAGAGGCTGATGCTGAATC 5866
QY 2690 AGTGCACCTGAGACAGGCTGAGTGTGTTCTTTCAGCTTGCCTTATGGGTAATCAT 2749
Db 5867 AGTGCACCTGAGACAGGCTGAGTGTGTTCTTTCAGCTTGCCTTATGGGTAATCAT 5926
QY 2750 GAGTCACGCACTGAGCCTTATGACATGAACCTGTGGAAGACATGCGCTGAAGCTTCTGT 2809
Db 5927 GAGTCACGCACTGAGCCTTATGACATGAACCTGTGGAAGACATGCGCTGAAGCTTCTGT 5986
QY 2810 GCGAGCTCATGAGAAACATTTGTCTATCTCCAGAGCCTGAGATTTGGTAAAGTGTCTATC 2869
Db 5987 GCGAGCTCATGAGAAACATTTGTCTATCTCCAGAGCCTGAGATTTGGTAAAGTGTCTATC 6046
QY 2870 TCACGCGCGCTGTGTGAGTGTCTGTCTGTGTATCTCGAGAGACAGACACTGGAAGA 2929
Db 6047 TCACGCGCGCTGTGTGAGTGTCTGTCTGTGTATCTCGAGAGACAGACACTGGAAGA 6106
QY 2930 GCTTGATCTTCAGCAATGCTGCTGAGTGAAGTGGGTTCTGCTGCTGTGCGAGGAGC 2989
Db 6107 GCTTGATCTTCAGCAATGCTGCTGAGTGAAGTGGGTTCTGCTGCTGTGCGAGGAGC 6166
QY 2990 TGAACCAAAAGAAAGTGTGTGAGAGACTGGGTTGAAGCATGTGACTGACTTCTG 3049
Db 6167 TGAACCAAAAGAAAGTGTGTGAGAGACTGGGTTGAAGCATGTGACTGACTTCTG 6226
QY 3050 ATTGCTGAGGACCTCTCTTGGCCCTTCTGCAACCGGATCTGACACTTAAACC 3109
Db 6227 ATTGCTGAGGACCTCTCTTGGCCCTTCTGCAACCGGATCTGACACTTAAACC 6286
QY 3110 TGGTCAGAAATTAATCTTCAAGTCCCAAGATGATGAGTGTGCTGAGCTTGTGCTGTC 3169
Db 6287 TGGTCAGAAATTAATCTTCAAGTCCCAAGATGATGAGTGTGCTGAGCTTGTGCTGTC 6346
QY 3170 CCAGCTCTTAATTAAGATTAATGG 3194
Db 6347 CCAGCTCTTAATTAAGATTAATGG 6371

```

RESULT 8  
 AAL47140  
 ID AAL47140 standard; DNA; 6939 BP.

XX AAL47140;  
 XX 20-AUG-2002 (first entry)  
 DE Pyrin domain containing protein NALP3/Py17 coding sequence.  
 XX Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;  
 XX antiarteriosclerotic; antiporiatic; antibacterial; virucide;  
 XX neuroprotective; antiarthritic; antirheumatic; antiaschematic;  
 XX nephroprotective; osteopathic; nootropic; intracellular signal transduction;  
 XX inflammation; Alzheimer's disease; infection; psoriasis; asthma;  
 XX arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;  
 XX osteoarthritis; glomerulonephritis; gene; ds.

```

OS Unidentified.
PN MO200240668-A2.
PD 23-MAY-2002.
PF 30-OCT-2001; 2001MO-EP012545.
PR 15-NOV-2000; 2000DE-01056687.
PR 30-NOV-2000; 2000DE-01059595.
XX (APOT-) APOTEC RES & DEV LTD.
XX PA
XX Tschoep J, Martimon F;
XX WPI, 2002-427093/45.
XX P-PSDB; AA017868.
XX
PT New DNA encoding protein with pyrin domain, useful for treating diseases
PT involving impaired signal transduction, particularly inflammation, also
PT proteins and antibodies.
XX
PS Claim 5; Fig 1; 116pp; German.
XX
CC The present invention relates the DNA and their encoded proteins, where
CC the proteins contain at least one PYD (pyrin) domain. These can be used
CC to treat diseases associated with impaired intracellular signal
CC transduction, particularly inflammation such as psoriasis,
CC arteriosclerosis, bacterial or viral infections (particularly meningitis
CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC the invention
XX
SQ Sequence 6939 BP; 1791 A; 1719 C; 1801 G; 1628 T; 0 U; 0 Other;
XX
Query Match 69.5%; Score 2803; DB 6; Length 6939;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2003; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 290 AAATTTCAGACTATGAGACAGAAAGGTGCGACAGCAGACAGACAGAAACAAAGAA 349
Db 3467 AAATTTCAGACTATGAGACAGAAAGGTGCGACAGCAGACAGACAGAAACAAAGAA 3526
QY 350 TTTCAAGACTATGAGACAGAAAGGTGCGACAGCAGACAGACAGAAACAAAGCAATG 409
Db 3527 TTTCAAGACTATGAGACAGAAAGGTGCGACAGCAGACAGACAGAAACAAAGCAATG 3586
QY 410 GAGGTGACACATGAGACCTTCAAGAGTCACTGATGACCAAAATTCGCTGAGAGAGATG 469
Db 3587 GAGGTGACACATGAGACCTTCAAGAGTCACTGATGACCAAAATTCGCTGAGAGAGATG 3646
QY 470 TACGTCGATGTTTGAAGAACTGCTGCTGAGTGGCCGGAATGCAAGTGTGCTG 529
Db 3647 TACGTCGATGTTTGAAGAACTGCTGCTGAGTGGCCGGAATGCAAGTGTGCTG 3706
QY 530 CTTTGTATCAGACCGGTGGGGCTTCCGGCTTCGACGAGTGTCTGACGAGAAAGTCA 589
Db 3707 CTTTGTATCAGACCGGTGGGGCTTCCGGCTTCGACGAGTGTCTGACGAGAAAGTCA 3766
QY 590 GAATTGGGAATTCGCTTACCCAGAAAGATCTGTGTGCTGGCCGGAAGTGAATCT 649
Db 3767 GAATTGGGAATTCGCTTACCCAGAAAGATCTGTGTGCTGGCCGGAAGTGAATCT 3826
QY 650 ACCAGGAATGTTCTCTAGCTCTTCTTCTCCCGTTAGAGATGACGGAAGAAAG 709
Db 3827 ACCAGGAATGTTCTCTAGCTCTTCTTCTCCCGTTAGAGATGACGGAAGAAAG 3886
QY 710 AGAGAGTGTCAAGAGTTCATCTCCAGAGAGTGGCCAGACTCCAGGCTCCGCTGACG 769
Db 3887 AGAGAGTGTCAAGAGTTCATCTCCAGAGAGTGGCCAGACTCCAGGCTCCGCTGACG 3946
QY 770 AGATCATGTCCTCCGACAGAAAGGCTGTTGTTCAATTCATGACGCTTGCATCCTGGGCT 829

```

Db 3947 AGATCATATCCGACCAAGAAAGGCTGTTGTTATCATTTGACGGTTTCATGACCTCGGGCT 4006  
Qy 830 CTGTCTTAACAAGTACCAAAAGCTCTGCAAAAGATCGGGCTGAGAGAGAGCTCCGTTCA 889  
Db 4007 CTGTCTTAACAAGTACCAAAAGCTCTGCAAAAGATCGGGCTGAGAGAGAGCTCCGTTCA 4066  
Qy 890 CCCATCATGAGAGTGTGCTGAGAGAGAGTCTGCTCCGAGAGCTCTGAGTCTGATCTGACCG 949  
Db 4067 CCTCATATGAGAGTGTGCTGAGAGAGTCTGCTCCGAGAGCTCTGAGTCTGATCTGACCG 4126  
Qy 950 TCAAGAGTGTGGGCAAGAGAGAGCTCAAGTCAAGAGTGTGTCTCCCGTTAACTGTTAG 1009  
Db 4127 TCAAGAGTGTGGGCAAGAGAGAGCTCAAGTCAAGAGTGTGTCTCCCGTTAACTGTTAG 4186  
Qy 1010 TTAAAGAAATCTCCGGGGAAAGAAAGATTCATCTGCTCTTGAAGCGGGATTTGTAGC 1069  
Db 4187 TTAAAGAAATCTCCGGGGAAAGAAAGATTCATCTGCTCTTGAAGCGGGATTTGTAGC 4246  
Qy 1070 ATCAAGAAACAAGGGTGTGGTGCATGATGAACAACGTGAGCTGTGACCAAGTGC 1129  
Db 4247 ATCAAGAAACAAGGGTGTGGTGCATGATGAACAACGTGAGCTGTGACCAAGTGC 4306  
Qy 1130 AGGTGCCCCGGGGCTCTCTCATCTGCTGAGCCCTGACAGCTGAGAGAGTGTGGGG 1189  
Db 4307 AGGTGCCCCGGGGCTCTCTCATCTGCTGAGCCCTGACAGCTGAGAGAGTGTGGGG 4366  
Qy 1190 AGAGGTGCCCCCTTCAACCAAGCTCAAGCTGACAGCGCTTTGTGTTATC 1249  
Db 4367 AGAGGTGCCCCCTTCAACCAAGCTCAAGCTGACAGCGCTTTGTGTTATC 4426  
Qy 1250 AGCTACCCCTGAGGGGTGGTGGGGCTGTCTCAATCTGAGAGAAAGATGTCTGA 1309  
Db 4427 AGCTACCCCTGAGGGGTGGTGGGGCTGTCTCAATCTGAGAGAAAGATGTCTGA 4486  
Qy 1310 AGCGTTTCTGCGCTATGAGTGTGAGAGAGTGTGAATAGAGAGTCACTGTTATGTTG 1369  
Db 4487 AGCGTTTCTGCGCTATGAGTGTGAGAGAGTGTGAATAGAGAGTCACTGTTATGTTG 4546  
Qy 1370 ACGACCTATAGTTCAGAGAGCTCGGGAGTCTGAGCTCCGTCTGTTTCAATGAACA 1429  
Db 4547 ACGACCTATAGTTCAGAGAGCTCGGGAGTCTGAGCTCCGTCTGTTTCAATGAACA 4606  
Qy 1430 TCCCTTCCGAGAGAGCACTGTGAGAGTACTACCTTCTCCACCTCACTGCTCCAGG 1489  
Db 4607 TCCCTTCCGAGAGAGCACTGTGAGAGTACTACCTTCTCCACCTCACTGCTCCAGG 4666  
Qy 1490 ACTTCTGTGCGGCTTGTACTACGTGTGAGAGGGCTGGAATCGAGCCAGCTCTGCGC 1549  
Db 4667 ACTTCTGTGCGGCTTGTACTACGTGTGAGAGGGCTGGAATCGAGCCAGCTCTGCGC 4726  
Qy 1550 CTCTGTATGTTGAGAGCAAAAGAGTCTCATGAGCTTTAAACAGGAGGCTTCCATATCC 1609  
Db 4727 CTCTGTATGTTGAGAGCAAAAGAGTCTCATGAGCTTTAAACAGGAGGCTTCCATATCC 4786  
Qy 1610 ACTGCTTTGATGAGAGGTTCTTGTGTGCTCTGAGAGAGAGCTAAGAGAGGCGAC 1669  
Db 4787 ACTGCTTTGATGAGAGGTTCTTGTGTGCTCTGAGAGAGAGCTAAGAGAGGCGAC 4846  
Qy 1670 TGGAGGTCTGCTGGGCTGTGCTCCGTTCCCTGAGGGGTGAGACAGAGCTTGTGACTGGG 1729  
Db 4847 TGGAGGTCTGCTGGGCTGTGCTCCGTTCCCTGAGGGGTGAGACAGAGCTTGTGACTGGG 4906  
Qy 1730 TCTCTCTGTGTGAGAGCTTAATGTCACCACTCCAGAGAGACCTCTGAGCGCTTCC 1789  
Db 4907 TCTCTCTGTGTGAGAGCTTAATGTCACCACTCCAGAGAGACCTCTGAGCGCTTCC 4966  
Qy 1790 ACTGCTTTTGAAGCTCAAGAGCAAAAGATTGTTGTGCTGTGCAATTAACAGCTTCCAG 1849  
Db 4967 ACTGCTTTTGAAGCTCAAGAGCAAAAGATTGTTGTGCTGTGCAATTAACAGCTTCCAG 5026  
Qy 1850 AAGTGTGCTTCCGATTAACAGAGACTGAGCTTGATAGCATCTTCTGTGCTCCAGC 1909  
|||||

Db 5027 AAGTGTGCTTCCGATTAACAGAGACTGAGCTTGATAGCATCTTCTGTGCTCCAGC 5086  
Qy 1910 ACTGTCCGTAATTTGCGGAAAAATTCGGGTGATGTCAAAAGGATTTTCCCAAGATAGT 1969  
Db 5087 ACTGTCCGTAATTTGCGGAAAAATTCGGGTGATGTCAAAAGGATTTTCCCAAGATAGT 5146  
Qy 1970 CCGCTGAGGCATGTCTGTGGTCCCTCTATAGATGCGGATTAAGACCTCATTTAGAGAGC 2029  
Db 5147 CCGCTGAGGCATGTCTGTGGTCCCTCTATAGATGCGGATTAAGACCTCATTTAGAGAGC 5206  
Qy 2030 AGTGGAGAAATTTCTGCTCCATGCTTGTGACCAACCAACCTGCGGAGCTGAGCTGG 2089  
Db 5207 AGTGGAGAAATTTCTGCTCCATGCTTGTGACCAACCAACCTGCGGAGCTGAGCTGG 5266  
Qy 2090 GCACAGATCTCTACAGAGCGGGCCATGAAAGACCTGTGTGCCAGCTGAGGATCCCA 2149  
Db 5267 GCACAGATCTCTACAGAGCGGGCCATGAAAGACCTGTGTGCCAGCTGAGGATCCCA 5326  
Qy 2150 CCTGCAAGATPACAGACCTGATGTTTGAAGATTAACCTGCTGTGTGAGAGACC 2209  
Db 5327 CCTGCAAGATPACAGACCTGATGTTTGAAGATTAACCTGCTGTGTGAGAGACC 5386  
Qy 2210 TCTGAGAAATGTCATGAGCCCAACCGTAACCTTAAGATCTTCACTTGGAGGACCAAC 2269  
Db 5387 TCTGAGAAATGTCATGAGCCCAACCGTAACCTTAAGATCTTCACTTGGAGGACCAAC 5446  
Qy 2270 TGAAGAGAGAGATGTAAGATGCGCTGTGAAGCTTTAAACACCCAAATGTTGTTGG 2329  
Db 5447 TGAAGAGAGAGATGTAAGATGCGCTGTGAAGCTTTAAACACCCAAATGTTGTTGG 5506  
Qy 2330 AGTCTTTGAGGCTGATTTGCTGTGATGACCCATGCTGTTAACCTGAAGATCTCCAAA 2389  
Db 5507 AGTCTTTGAGGCTGATTTGCTGTGATGACCCATGCTGTTAACCTGAAGATCTCCAAA 5566  
Qy 2390 TCTTTAGACCTCCCAAGCTTGAATCTGTAGCTTGTGAGAGAAACAAGTGAACAAC 2449  
Db 5567 TCTTTAGACCTCCCAAGCTTGAATCTGTAGCTTGTGAGAGAAACAAGTGAACAAC 5626  
Qy 2450 AGGAGATTAATGCCCTGATGATGACCTGAGAGTCTCCAGTGGCCCTGAGAGAGCTGA 2509  
Db 5627 AGGAGATTAATGCCCTGATGATGACCTGAGAGTCTCCAGTGGCCCTGAGAGAGCTGA 5686  
Qy 2510 TACTGAGAGACTGTGAGATCAACAGCAGGTTTCCAGAGTCTGAGCTCAGCCCTGCTCA 2569  
Db 5687 TACTGAGAGACTGTGAGATCAACAGCAGGTTTCCAGAGTCTGAGCTCAGCCCTGCTCA 5746  
Qy 2570 GCAACCGGAGCTTGAACAACCTGTGCTTATCAACAACAGCTGGGAAACGAGCTGTAA 2629  
Db 5747 GCAACCGGAGCTTGAACAACCTGTGCTTATCAACAACAGCTGGGAAACGAGCTGTAA 5806  
Qy 2630 ATCTAATGTGTGATTCATAGAGCTTCCCACTGTATCTGCAAGAGCTGATGCTGAATC 2689  
Db 5807 ATCTAATGTGTGATTCATAGAGCTTCCCACTGTATCTGCAAGAGCTGATGCTGAATC 5866  
Qy 2690 AGTGCACCTGAGACAGGCTGTGCTGTGTTTCTTGTGACTTGTGAGCTTATGGGTAACTCAT 2749  
Db 5867 AGTGCACCTGAGACAGGCTGTGCTGTGTTTCTTGTGACTTGTGAGCTTATGGGTAACTCAT 5926  
Qy 2750 GGTGACGACCTTGAAGCTTATGATGAACCTGTGGAAGACAATGCGTGAAGCTTCTGT 2809  
Db 5927 GGTGACGACCTTGAAGCTTATGATGAACCTGTGGAAGACAATGCGTGAAGCTTCTGT 5986  
Qy 2810 GCGAGGTCAATGAGAGAACATCTTGTGATCTCCAGAGACTGTGAGATTGTAAAGTGTATC 2869  
Db 5987 GCGAGGTCAATGAGAGAACATCTTGTGATCTCCAGAGACTGTGAGATTGTAAAGTGTATC 6046  
Qy 2870 TCACCGCGGCTGTGTGAGAGTCTGTCTGTGTGATCTGAGAGAGAGACACTGAAGA 2929  
Db 6047 TCACCGCGGCTGTGTGAGAGTCTGTCTGTGTGATCTGAGAGAGAGACACTGAAGA 6106  
Qy 2930 GCTGTGATTTCAACGAGCAATGCCCTGTGGGTGAACGTTGGGTTGTGCTGCTGTGAGAGAC 2989  
Db 6107 GCTGTGATTTCAACGAGCAATGCCCTGTGGGTGAACGTTGGGTTGTGCTGCTGTGAGAGAC 6166  
|||||

QY 2990 TGAAGCAAAAGAGAGCTGTTCTGACGAGACTGGGTTGAAGGACGATGACCTTTCG 3049  
DB 6167 TGAAGCAAAAGAGAGCTGTTCTGACGAGACTGGGTTGAAGGACGATGACCTTTCG 6226  
QY 3050 ATTGCTGTGAGGACGACTCTCTGCTTGGCCCTTCTGCAACCGGATCTGACCACTTAAAC 3109  
DB 6227 ATTGCTGTGAGGACGACTCTCTGCTTGGCCCTTCTGCAACCGGATCTGACCACTTAAAC 6286  
QY 3110 TGGTCAGAAATTAATCTCACTGCCAAGAGATGAAAGCTGTGCTGGCCTTTCCTGTC 3169  
DB 6287 TGGTCAGAAATTAATCTCACTGCCAAGAGATGAAAGCTGTGCTGGCCTTTCCTGTC 6346  
QY 3170 CCAGCTCTACTTACGATTAATGG 3194  
DB 6347 CCAGCTCTACTTACGATTAATGG 6371

RESULT 9  
ABX97181  
ID ABX97181 standard; cDNA, 3226 BP.  
XX  
AC ABX97181;  
XX  
DT 20-MAY-2003 (first entry)  
XX  
DE Human NOV125b cDNA.  
XX  
KM NOVA; cytosolic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;  
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;  
XX human; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200272757-A2.  
XX  
PD 19-SEP-2002.  
XX  
PF 08-MAR-2002; 2002WO-US006908.  
XX  
PR 08-MAR-2001; 2001US-0274101P.  
PR 08-MAR-2001; 2001US-0274194P.  
PR 08-MAR-2001; 2001US-0274281P.  
PR 08-MAR-2001; 2001US-0274322P.  
PR 09-MAR-2001; 2001US-0274849P.  
PR 12-MAR-2001; 2001US-0275235P.  
PR 13-MAR-2001; 2001US-0275578P.  
PR 13-MAR-2001; 2001US-0275579P.  
PR 13-MAR-2001; 2001US-0275601P.  
PR 14-MAR-2001; 2001US-0276000P.  
PR 16-MAR-2001; 2001US-0276767P.  
PR 19-MAR-2001; 2001US-0276994P.  
PR 20-MAR-2001; 2001US-0277239P.  
PR 20-MAR-2001; 2001US-0277321P.  
PR 20-MAR-2001; 2001US-0277327P.  
PR 21-MAR-2001; 2001US-0277791P.  
PR 22-MAR-2001; 2001US-0277833P.  
PR 23-MAR-2001; 2001US-0278152P.  
PR 26-MAR-2001; 2001US-0278894P.  
PR 27-MAR-2001; 2001US-0278999P.  
PR 27-MAR-2001; 2001US-0279035P.  
PR 28-MAR-2001; 2001US-0279344P.  
PR 30-MAR-2001; 2001US-0277338P.  
PR 30-MAR-2001; 2001US-0279995P.  
PR 30-MAR-2001; 2001US-0280233P.  
PR 02-APR-2001; 2001US-0280822P.  
PR 02-APR-2001; 2001US-0280822P.  
PR 02-APR-2001; 2001US-0280900P.  
PR 04-APR-2001; 2001US-0281194P.  
PR 13-APR-2001; 2001US-0283675P.  
PR 30-APR-2001; 2001US-0287424P.  
PR 02-MAY-2001; 2001US-0288066P.  
PR 03-MAY-2001; 2001US-0288342P.

PR 03-MAY-2001; 2001US-0288528P.  
PR 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291099P.  
PR 16-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
PR 31-MAY-2001; 2001US-0294889P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 18-JUN-2001; 2001US-0299027P.  
PR 19-JUN-2001; 2001US-0299303P.  
PR 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
PR 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
PR 10-SEP-2001; 2001US-0318462P.  
PR 12-SEP-2001; 2001US-0318770P.  
PR 27-SEP-2001; 2001US-0325430P.  
PR 27-SEP-2001; 2001US-0325681P.  
PR 18-OCT-2001; 2001US-0330380P.  
PR 31-OCT-2001; 2001US-0333011P.  
PR 14-NOV-2001; 2001US-0333172P.  
PR 14-NOV-2001; 2001US-0333271P.  
PR 14-NOV-2001; 2001US-0333272P.  
PR 14-NOV-2001; 2001US-0333184P.  
PR 14-NOV-2001; 2001US-0333272P.  
PR 14-NOV-2001; 2001US-0333272P.  
PR 21-NOV-2001; 2001US-0333094P.  
PR 03-DEC-2001; 2001US-0337426P.  
PR 03-DEC-2001; 2001US-0338092P.  
PR 04-DEC-2001; 2001US-0337185P.  
PR 03-JAN-2002; 2002US-0345705P.  
PR 07-MAR-2002; 2002US-00092900.  
XX  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Padigaru M, Spytek KA, Shenoy SG, Taupier RU, Pena CEA, Li L;  
PI Zernusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;  
PI Patirajan M, Gangoli E, Vermet CAM, Guo X, Tchernev V;  
PI Fernandes ER, Casman ST, Malvankar UM, Gerlach V, Liu Y, Anderson D;  
PI Spaderina SK, Catterton E, Burgess C, Lette M, Zhong H, Alsdbrook JP;  
PI Lepley DM, Rieger DK;  
XX  
XX WPI: 2002-723332/78.  
XX DR P-PsDB; ABU65214.  
XX  
XX NOVA polypeptides and polynucleotides, useful for preventing or treating  
PT a disorder associated with aberrant NOVA expression or activity e.g.,  
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial  
PT asthma.  
XX  
XX Claim 13; Page 467; 1103pp; English.  
XX  
XX This invention describes novel human NOVA polypeptides which have  
CC cytosolic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive  
CC activity. Pharmaceutical compositions comprising the NOVA proteins or  
CC nucleic acid molecules or NOVA antibodies are useful for preventing or  
CC treating a disorder associated with aberrant NOVA expression or activity  
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial  
CC asthma. The products of the invention can be used for gene therapy or in  
CC a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by  
CC the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-  
CC ABX97593. ABX97008-ABX97185 encode the NOVA proteins described in  
CC ABU65041-ABU65218  
XX  
SQ Sequence 3226 BP; 738 A; 856 C; 914 G; 718 T; 0 U; 0 Other;

Query Match 69.4%; Score 2801; DB 6; Length 3226;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2501; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 406 CATGAGGTGACACATGAGGATACAGAGTACGATGATGACCAATTCGCTGAGGAGAG 465  
DB 249 CATGAGGTGACACATGAGGATACAGAGTACGATGATGACCAATTCGCTGAGGAGAG 308  
QY 466 GATGACGTGCTAGTTTGAAGAACACTGCTCTGACTGCGCGGAAATGCAACGTTGCT 525

Db 309 GATGACGTCTAGTTTGAACACATGCTGCTGACCTGGCCGGAATCAACGTGGCT 368  
Qy 526 GGTCTTTTGAATTCAGACCGGTGGGCTTCCGGCTCGCAGCGGTGTTTCACAGGAAG 585  
Db 369 GGTCTTTTGAATTCAGACCGGTGGGCTTCCGGCTCGCAGCGGTGTTTCACAGGAAG 428  
Qy 586 TCAGGAATTTGGGAATCCGCTCTAGCCAGAGGATCTGTGCTGGGCCAAGTGA 645  
Db 429 TCAGGAATTTGGGAATCCGCTCTAGCCAGAGGATCTGTGCTGGGCCAAGTGA 488  
Qy 646 CTTCACAGGGAATGTTCTCTAGCTTTCTTCTCCCGTTAAGAGATCAAGCGGAAG 705  
Db 489 CTTCACAGGGAATGTTCTCTAGCTTTCTTCTCCCGTTAAGAGATCAAGCGGAAG 548  
Qy 706 AAGGAGCAGTGTCAAGATTCACAGGAGTGGCCAGACTCCAGGCTCCGGTG 765  
Db 549 AAGGAGCAGTGTCAAGATTCACAGGAGTGGCCAGACTCCAGGCTCCGGTG 608  
Qy 766 ACGGAGATCATGTCCCGACAGAAAGCTGTGTTCAATCATTTGACGTTTCAGATGACCTG 825  
Db 609 ACGGAGATCATGTCCCGACAGAAAGCTGTGTTCAATCATTTGACGTTTCAGATGACCTG 668  
Qy 826 GGTCTGTCTCAACATGACACAAAGCTTTCAGAAAGCTGGCTGAGAGAGCCTTCG 885  
Db 669 GGTCTGTCTCAACATGACACAAAGCTTTCAGAAAGCTGGCTGAGAGAGCCTTCG 728  
Qy 886 TTCAACCTCATACCAAGTCTGTGAGGAAGTCCGTGCTCCCTGAGTCCCTTCATATGCTC 945  
Db 729 TTCAACCTCATACCAAGTCTGTGAGGAAGTCCGTGCTCCCTGAGTCCCTTCATATGCTC 788  
Qy 946 ACCGTCAAGAGACGTGGGACAGAGAGCTCAAGTCAGAGTCTGTTCTCCCGTTACCTG 1005  
Db 789 ACCGTCAAGAGACGTGGGACAGAGAGCTCAAGTCAGAGTCTGTTCTCCCGTTACCTG 848  
Qy 1006 TTAGTTAAGGAATCTCCGGGAAACAAAGAAATCCAATTGCTCTTGAAGCGCGGAATGGT 1065  
Db 849 TTAGTTAAGGAATCTCCGGGAAACAAAGAAATCCAATTGCTCTTGAAGCGCGGAATGGT 908  
Qy 1066 GAGGATCAAGAGACAAAGGTTGCTGTCGATCAATGAACAACGTCGAGCTTCGACGAC 1125  
Db 909 GAGGATCAAGAGACAAAGGTTGCTGTCGATCAATGAACAACGTCGAGCTTCGACGAC 968  
Qy 1126 TGCCAGGTCCCGCGCTGCTCTCATCTGCTGAGCTTCGACGTCAGGAGCTGGTG 1185  
Db 969 TGCCAGGTCCCGCGCTGCTCTCATCTGCTGAGCTTCGACGTCAGGAGCTGGTG 1028  
Qy 1186 GGGGAGAGCTGCGCCCTTCAACCAACGCTCACAGGCTTCGACGCGCTTTTGTGTTT 1245  
Db 1029 GGGGAGAGCTGCGCCCTTCAACCAACGCTCACAGGCTTCGACGCGCTTTTGTGTTT 1088  
Qy 1246 CATCAGCTCACCCCTCGAGGCGTGGTCCGGGCTGTCTCATCTGAGGAGAAAGTTGTC 1305  
Db 1089 CATCAGCTCACCCCTCGAGGCGTGGTCCGGGCTGTCTCATCTGAGGAGAAAGTTGTC 1148  
Qy 1306 CTGAAGCGCTTCTGCGGTATGCTGTGAGGAGTGTGAATAGGAAGTCAGTGTTCAT 1365  
Db 1149 CTGAAGCGCTTCTGCGGTATGCTGTGAGGAGTGTGAATAGGAAGTCAGTGTTCAT 1208  
Qy 1366 GGTGACGACTTCATGTTCAAGACTCGGGAGTCTGAGCTCCGTCTGTGTTCAATG 1425  
Db 1209 GGTGACGACTTCATGTTCAAGACTCGGGAGTCTGAGCTCCGTCTGTGTTCAATG 1268  
Qy 1426 AACATCTTCTCCAGACAGCACTGTGAGGAGTCAACCTTCTTCAACCTGATGTC 1485  
Db 1269 AACATCTTCTCCAGACAGCACTGTGAGGAGTCAACCTTCTTCAACCTGATGTC 1328  
Qy 1486 CAGGACTTCTGTGCGCTTGTACTACGTGTTAAGGGGCTGGAATCAGACGACTTC 1545  
Db 1329 CAGGACTTCTGTGCGCTTGTACTACGTGTTAAGGGGCTGGAATCAGACGACTTC 1388  
Qy 1546 TGCCCTCTGTACGTTGAGAGACAAAGAGTCAATGAGCTTAAACAGGAGGCTTCAT 1605

Db 1389 TGCCCTCTGTACGTTGAGAAACAAAGAGTCCATGAGCTTAAACAGCAGGCTTCAT 1448  
Qy 1606 ATCCACTCGCTTTGGATGAAGCGTTTCTGTTGGCTCGTGAGCGAAGAGTAAGAG 1665  
Db 1449 ATCCACTCGCTTTGGATGAAGCGTTTCTGTTGGCTCGTGAGCGAAGAGTAAGAG 1508  
Qy 1666 CCACTGAGAGTCTGCTGGGCTGTCCGTTCCCTCGGAGGTGAAGCAGAAAGCTTCGAC 1725  
Db 1509 CCACTGAGAGTCTGCTGGGCTGTCCGTTCCCTCGGAGGTGAAGCAGAAAGCTTCGAC 1568  
Qy 1726 TGGGTCTCTGTGTGGGTCAAGACCTTAATGCCACCAACCCAGAGACACCTTGGAGGCC 1785  
Db 1569 TGGGTCTCTGTGTGGGTCAAGACCTTAATGCCACCAACCCAGAGACACCTTGGAGGCC 1628  
Qy 1786 TTCACTGTCTTTTCAGACTCAAGACAAAGAGTTTGTGCTTGGCATTTAAACAGCTTC 1845  
Db 1629 TTCACTGTCTTTTCAGACTCAAGACAAAGAGTTTGTGCTTGGCATTTAAACAGCTTC 1688  
Qy 1846 CAAGAGTGTGGCTTCCGATTTAACAGAACCTGAGCTTGATAGCATTTCTCTGCTTC 1905  
Db 1689 CAAGAGTGTGGCTTCCGATTTAACAGAACCTGAGCTTGATAGCATTTCTCTGCTTC 1748  
Qy 1906 CAGCACTGTCCGTATTTGGCGGAAATTCGGGTGATGTCAAGGAGATCTTCCAAAGAT 1965  
Db 1749 CAGCACTGTCCGTATTTGGCGGAAATTCGGGTGATGTCAAGGAGATCTTCCAAAGAT 1808  
Qy 1966 GAGTCCGCTGAGGAGATGCTGTGTCCTCTATGATGCGGATTAAGACCTCATTTGAG 2025  
Db 1809 GAGTCCGCTGAGGAGATGCTGTGTCCTCTATGATGCGGATTAAGACCTCATTTGAG 1868  
Qy 2026 GAGCAGTGGGAAGATTTCTGCTCCATGCTTGGACCCACCCACACTGCGGAGCTGGAC 2085  
Db 1869 GAGCAGTGGGAAGATTTCTGCTCCATGCTTGGACCCACCCACACTGCGGAGCTGGAC 1928  
Qy 2086 CTGGGACAGCAATCTCTGACAGAGCGGCCATGAAGACCTGTGTGCCAAGCTGAGCAT 2145  
Db 1929 CTGGGACAGCAATCTCTGACAGAGCGGCCATGAAGACCTGTGTGCCAAGCTGAGCAT 1988  
Qy 2146 CCCACTGCAAGATACAGACCCCTGATGTTAGAAATGACAGATTAACCCCTGGTGTGAG 2205  
Db 1989 CCCACTGCAAGATACAGACCCCTGATGTTAGAAATGACAGATTAACCCCTGGTGTGAG 2048  
Qy 2206 CACCTTGAAGAAATGTCATGTGGCCAAACCTGAACCTTCATCACTTGGAGGAC 2265  
Db 2049 CACCTTGAAGAAATGTCATGTGGCCAAACCTGAACCTTCATCACTTGGAGGAC 2108  
Qy 2266 CACTGAAGAGAAAGATGTAAGATGGCTGTGAACCTTTAAACACCCAAAATGTTTG 2325  
Db 2109 CACTGAAGAGAAAGATGTAAGATGGCTGTGAACCTTTAAACACCCAAAATGTTTG 2168  
Qy 2326 TTGAGTCTTTGAGGCTGGAATGCTGTGATTTGACCCATGCTGTACTGAAGATCTCC 2385  
Db 2169 TTGAGTCTTTGAGGCTGGAATGCTGTGATTTGACCCATGCTGTACTGAAGATCTCC 2228  
Qy 2386 CAAATCTTACGACCTTCCCGACGCTGAATTTCTGAGCTTCGAGGAAACAAAGTGA 2445  
Db 2229 CAAATCTTACGACCTTCCCGACGCTGAATTTCTGAGCTTCGAGGAAACAAAGTGA 2288  
Qy 2446 GACCAAGGAGTAATGCTCTCAGTATGCTTGAAGATCTTCCAGTGGCGCTTCAGAG 2505  
Db 2289 GACCAAGGAGTAATGCTCTCAGTATGCTTGAAGATCTTCCAGTGGCGCTTCAGAG 2348  
Qy 2506 CTGATCTGAGGAGCTGTGGATCAAGCAGCGGTTGCCAGAGTCTGGCTCAGGCTTC 2565  
Db 2349 CTGATCTGAGGAGCTGTGGATCAAGCAGCGGTTGCCAGAGTCTGGCTCAGGCTTC 2408  
Qy 2566 GTCAGCAACCGGAGCTTGAACAACCTGTGCTATTCACAACAAGCTTGGGAAAGAAAGT 2625  
Db 2409 GTCAGCAACCGGAGCTTGAACAACCTGTGCTATTCACAACAAGCTTGGGAAAGAAAGT 2468  
Qy 2626 GTAAATCTACTGTGTGCATCAATGAGGCTTCCCATGTATGTGCAAGAGCTGATGCTG 2685  
Db 2469 GTAAATCTACTGTGTGCATCAATGAGGCTTCCCATGTATGTGCAAGAGCTGATGCTG 2528

QY 2686 AATCAGTCGCACTTGAGACAGCGCTGGCTGTTTCTTGCACTTGCGCTTAAGGGTAAC 2745  
DB 2529 AATCAGTCGCACTTGAGACAGCGCTGGCTGTTTCTTGCACTTGCGCTTAAGGGTAAC 2588  
QY 2746 TCATGCTGACGCACTTGAGACAGCGCTTGACATGAACCTTGAGAAACAAATGGCGTAAGCTT 2805  
DB 2589 TCATGCTGACGCACTTGAGACAGCGCTTGACATGAACCTTGAGAAACAAATGGCGTAAGCTT 2648  
QY 2806 CTGTGAGAGTCATGAGAGAACCATTTGTCTCATCTCCAGACCTTGAGATTGGTAAGTGT 2865  
DB 2649 CTGTGAGAGTCATGAGAGAACCATTTGTCTCATCTCCAGACCTTGAGATTGGTAAGTGT 2708  
QY 2866 CATCTCACCGCGCGCTGTGAGAGTCGTCTGTGTATCTCGAGAGACAGACCTTG 2925  
DB 2709 CATCTCACCGCGCGCTGTGAGAGTCGTCTGTGTATCTCGAGAGACAGACCTTG 2768  
QY 2926 AAGAGCTTGATCTCAACGACAAATGCTGAGTGAAGTGTGCTGCTGCTGCTGAG 2985  
DB 2769 AAGAGCTTGATCTCAACGACAAATGCTGAGTGAAGTGTGCTGCTGCTGCTGAG 2828  
QY 2986 GGAATGAGCAAAAGAACAGTGTCTGACGAGACTCGGCTTGAGAGCATGTGACTGACT 3045  
DB 2829 GGAATGAGCAAAAGAACAGTGTCTGACGAGACTCGGCTTGAGAGCATGTGACTGACT 2888  
QY 3046 TCTGATGCTGAGAGCACTCTCTTGCGCTTCTTGACACCGCATCTGACCACTCTA 3105  
DB 2889 TCTGATGCTGAGAGCACTCTCTTGCGCTTCTTGACACCGCATCTGACCACTCTA 2948  
QY 3106 AACCTGTCGCAAAATTAATCTTCAAGTCCCAAGAAATGATGATGCTGCTGCTTGGC 3165  
DB 2949 AACCTGTCGCAAAATTAATCTTCAAGTCCCAAGAAATGATGATGCTGCTGCTTGGC 3008  
QY 3166 TGTCCAGCTCTAATCTTCAAGTAAATTTGGCTGTGAAATGACAGTACCTGTGCAATA 3225  
DB 3009 TGTCCAGCTCTAATCTTCAAGTAAATTTGGCTGTGAAATGACAGTACCTGTGCAATA 3068  
QY 3226 AGGAGCTGCTGAGAGAGTGAAGTCACTCAAGCCCGAGTGTATTAAGCTGACTTGG 3285  
DB 3069 AGGAGCTGCTGAGAGAGTGAAGTCACTCAAGCCCGAGTGTATTAAGCTGACTTGG 3128  
QY 3286 CATCTTTTGTAGATGATGACCG 3308  
DB 3129 CATCTTTTGTAGATGATGACCG 3151

RESULT 10  
ABK48628  
ID ABK48628 standard; cDNA; 3900 BP.  
XX  
AC ABK48628;  
XX  
DT 13-AUG-2002 (first entry)  
XX  
DE Human MATER cDNA.  
XX  
KW Human; gene; ss; contraceptive; antiinfertility; MATER;  
KW maternal antigen that embryos require; MATER null phenotype; oocyte;  
KW early embryonic survival; premature ovarian failure; POF;  
KW autoimmune infertility; chromosome 19; gene therapy; fertility.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..3603  
XX FT /\*tag= a  
XX FT /product= "Human MATER"  
XX  
XX WO200232955-A1.  
XX 25-APR-2002.  
XX  
XX 04-APR-2001; 2001WO-US010981.  
XX

XX  
PR 18-OCT-2000; 2000US-0241510P.  
XX  
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Nelson LM, Tong Z;  
XX  
XX WPI: 2002-454595/48.  
XX P-PDB; AAU79526.  
XX  
PT New isolated human Maternal Antigen That Embryos Require protein and  
PT polynucleotide, useful in diagnosing or treating fertility or reduced  
PT fertility, or as a contraceptive.  
XX  
XX Claim 11: Page 80-85; 93pp; English.  
XX  
XX The invention discloses an isolated human MATER (Maternal Antigen That  
XX Embryos Require) protein, which can complement a MATER null phenotype in  
XX which zygotes arising from the oocyte do not progress beyond the two-cell  
XX stage. MATER is required for early embryonic survival and abnormal levels  
XX of the protein can lead to premature ovarian failure (POF) and can be  
XX caused by under or over expression of MATER or an autoimmune response to  
XX MATER. MATER is a single-copy maternal effect gene found on chromosome  
XX 19. The MATER protein and polynucleotide, by gene therapy, are useful in  
XX diagnosing or treating fertility and reduced fertility. In particular,  
XX the MATER protein is useful as a contraceptive agent, or for influencing  
XX (either inhibiting or enhancing) fertility and can be used to detect a  
XX predisposition to infertility or reduced fertility, or for presymptomatic  
XX screening of an individual for infertility/reduced fertility. The protein  
XX and polynucleotide are also useful for detecting an excess or deficiency,  
XX or genetic mutation, of the MATER protein in a mammalian subject (e.g. a  
XX human or a mouse) or for screening for a compound useful in influencing  
XX MATER-mediated fertility. The sequence presented is the human MATER cDNA.  
XX Note: Nucleotides 1-75 were deduced, by the inventors, by comparison to  
XX published human genomic DNA sequences, while the remainder of the  
XX sequence was determined by direct cloning of human ovarian cDNAs  
XX

Sequence: 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;  
Query Match 67.2%; Score 2713; DB 6; Length 3900;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3013; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 290 AAATTTCAAGCTATGGAACAAGAGTGGCCACAGCAGCAGAGACAGAGAAACAAGAAA 349  
DB 566 AAATTTCAAGCTATGGAACAAGAGTGGCCACAGCAGCAGAGACAGAGAAACAAGAAA 625  
QY 350 TTTCAAGCTATGGAACAAGAGTGGCCACAGCAGCAGAGACAGAGAAACAAGACAATG 409  
DB 626 TTTCAAGCTATGGAACAAGAGTGGCCACAGCAGCAGAGACAGAGAAACAAGACAATG 685  
QY 410 GAGGTGACATGAGGACATCAAGAGTCAAGTATGACCAAAATTCGCTGAGAGAGATG 469  
DB 686 GAGGTGACATGAGGACATCAAGAGTCAAGTATGACCAAAATTCGCTGAGAGAGATG 745  
QY 470 TACGTGATGTTTGAAGAACTGCTGCTGAGCTGGCGGAAATGCAAACTTTGCTGCTG 529  
DB 746 TACGTGATGTTTGAAGAACTGCTGCTGAGCTGGCGGAAATGCAAACTTTGCTGCTG 805  
QY 530 CTTTGTATTCAGACCGGTGGGCTTCCGCTCGACAGTGTGTTGACAGGAAATGCTAG 589  
DB 806 CTTTGTATTCAGACCGGTGGGCTTCCGCTCGACAGTGTGTTGACAGGAAATGCTAG 865  
QY 590 GAATTTGGAATAATCGCTTACCCAGAGAGATGCTGCTGAGCGCAAGGTGACTCT 649  
DB 866 GAATTTGGAATAATCGCTTACCCAGAGAGATGCTGCTGAGCGCAAGGTGACTCT 925  
QY 650 ACCAGGGAATGTTCTCTACGCTTTCTTCTCCCGTTAGAGAGATGACGGGAAAGAG 709  
DB 926 ACCAGGGAATGTTCTCTACGCTTTCTTCTCCCGTTAGAGAGATGACGGGAAAGAG 985  
QY 710 AGAGAGTGCACAGAGTTCATCTCCAGGAGTGGCCAGACTCCAGACTCCGGTGAAGG 769

Db 986 AGACAGGTGTCACAGAGTTCTCTCCAGGAGTGGCCAGACTCCCGAGCTCCGGTGAACG 1045  
Qy 770 AGATCATGTCCCGACCAAGAAAGCGTGTTCATCATTTGACGTTTTGATGACCTGGGCT 829  
Db 1046 AGATCATGTCCCGACCAAGAAAGCGTGTTCATCATTTGACGTTTTGATGACCTGGGCT 1105  
Qy 830 CTGTCTCAACAAATGACAAAGCTCTGCAAAAGACTGGGCTGAGAGAGAGCTCCGTTCA 889  
Db 1106 CTGTCTCAACAAATGACAAAGCTCTGCAAAAGACTGGGCTGAGAGAGAGCTCCGTTCA 1165  
Qy 890 CCCCTCATGCGAGTCTGTGAGAGAGCTCTGCTCCCTGATCCTTCTGATTCGTCAACG 949  
Db 1166 CCCCTCATGCGAGTCTGTGAGAGAGCTCTGCTCCCTGATCCTTCTGATTCGTCAACG 1225  
Qy 950 TCAGAGAGTGGGACAGAGAGCTCAAGTCAGAGGTGTGTCTCCCGTTACCTGTTAG 1009  
Db 1226 TCAGAGAGTGGGACAGAGAGCTCAAGTCAGAGGTGTGTCTCCCGTTACCTGTTAG 1285  
Qy 1010 TTAGAGGATCTCCGGGGAAACAAAGATCACTTGCTCTTGAAGCGCGGAAATTGGTAGC 1069  
Db 1286 TTAGAGGATCTCCGGGGAAACAAAGATCACTTGCTCTTGAAGCGCGGAAATTGGTAGC 1345  
Qy 1070 ATCAGAGACACAAAGGTTGGGTGCGATCATGAAACACCTGAGCTGCTCCGACAGTCC 1129  
Db 1346 ATCAGAGACACAAAGGTTGGGTGCGATCATCMAACCGTGAAGCTGCTCCGACAGTCC 1405  
Qy 1130 AGGTGCGCGCGGTGGGCTCTCTCATCTGCGGCTGACAGCTGACAGAGCGTGGTGGGG 1189  
Db 1406 AGGTGCGCGCGGTGGGCTCTCTCATCTGCGGCTGACAGAGCTGAGAGCGTGGTGGGG 1465  
Qy 1190 AGAGCGTGGCCCCCTTCAACAAACGCTCAGAGGCTGACAGCGCTTTTGTGTTATC 1249  
Db 1466 AGAGCGTGGCCCCCTTCAACAAACGCTCAGAGGCTGACAGCGCTTTTGTGTTATC 1525  
Qy 1250 AGCTCAACCCCTGAGAGCGTGGTCCGCGCTGTCTCAATCTGAGAGAAAGTGTCTCGA 1309  
Db 1526 AGCTCAACCCCTGAGAGCGTGGTCCGCGCTGTCTCAATCTGAGAGAAAGTGTCTCGA 1585  
Qy 1310 AGCGCTTCTGCGATGAGCTGTGAGAGGAGTGTGAATAGAGAGTCACTGTTTGAATG 1369  
Db 1586 AGCGCTTCTGCGATGAGCTGTGAGAGGAGTGTGAATAGAGAGTCACTGTTTGAATG 1645  
Qy 1370 ACGACCTCATGATTCAGAGACTCGGAGAGTCTGAGACTCGTCTGTTTCAATGAA 1429  
Db 1646 ACGACCTCATGATTCAGAGACTCGGAGAGTCTGAGACTCGTCTGTTTCAATGAA 1705  
Qy 1430 TCTCTTCTCCAGACAGCCTGTGAGAGTACTACACTTCTTCAACTCACTCTCCAG 1489  
Db 1706 TCTCTTCTCCAGACAGCCTGTGAGAGTACTACACTTCTTCAACTCACTCTCCAG 1765  
Qy 1490 ACTTCTGCGCGCTGTACTACAGTGTAGAGGCGTGGAAATGAGGCACTCTCTG 1549  
Db 1766 ACTTCTGCGCGCTGTACTACAGTGTAGAGGCGTGGAAATGAGGCACTCTCTG 1825  
Qy 1550 CTCTGTACGTTGAGAGCAAAAGAGTCCATGAGACTTAAACAGGCAAGCTTCCATATC 1609  
Db 1826 CTCTGTACGTTGAGAGCAAAAGAGTCCATGAGACTTAAACAGGCAAGCTTCCATATC 1885  
Qy 1610 ACTGCTTTGATGAGAGGCTTCTTGTGTGCTGTGAGGAGAGAGCTAAGAGGCGAC 1669  
Db 1886 ACTGCTTTGATGAGAGGCTTCTTGTGTGCTGTGAGGAGAGAGCTAAGAGGCGAC 1945  
Qy 1670 TGAAGGTCTGTGAGGCTGTCCCGTTCCCTGAGGAGTGAAGAGAAAGCTTCTGCACTGG 1729  
Db 1946 TGAAGGTCTGTGAGGCTGTCCCGTTCCCGTTCCCTGAGGAGTGAAGAGAAAGCTTCTGCACTGG 2005  
Qy 1730 TCTCTGTGTTGAGTCAAGAGCTTAATGCAACCCGAGAGACACCTGAGAGCTTCC 1789  
Db 2006 TCTCTGTGTTGAGTCAAGAGCTTAATGCAACCCGAGAGACACCTGAGAGCTTCC 2065  
Qy 1790 ACTGTCTTTTGAAGCTCAAGAGCAAAAGTGTGTTGCTTGGCATTTAAACAGCTTCAAG 1849  
Db 2066 ACTGTCTTTTGAAGCTCAAGAGCAAAAGTGTGTTGCTTGGCATTTAAACAGCTTCAAG 2125

Qy 1850 AAGTGTGGCTTCCGATTAACAGAACCTGAGCTTGATAGATCTTCTTGTGCTTCGAGC 1909  
Db 2126 AAGTGTGGCTTCCGATTAACAGAACCTGAGCTTGATAGATCTTCTTGTGCTTCGAGC 2185  
Qy 1910 ACTGTCCGATTTTGGGAAAAATTTGGGTGATGTCAAAAGGATTTTCCCAAGATGAGT 1969  
Db 2186 ACTGTCCGATTTTGGGAAAAATTTGGGTGATGTCAAAAGGATTTTCCCAAGATGAGT 2245  
Qy 1970 CCGCTGAGGAGATGTCGTGTGCTCCCTGATAGATGCGGGATTAAGACCTCATTTAGAGAGC 2029  
Db 2246 CCGCTGAGGAGATGTCGTGTGCTCCCTGATAGATGCGGGATTAAGACCTCATTTAGAGAGC 2305  
Qy 2030 AGTGGGAAATTTCTGCTCCATGCTTGGCAACCCACCAACCTGCGAGCTGAGACCTGG 2089  
Db 2306 AGTGGGAAATTTCTGCTCCATGCTTGGCAACCCACCAACCTGCGAGCTGAGACCTGG 2365  
Qy 2090 GCAGCAGATCTTGAACAGAGCGGCGCATGAAAGACCTGTGTGCCAAGCTGAGGATCCCA 2149  
Db 2366 GCAGCAGATCTTGAACAGAGCGGCGCATGAAAGACCTGTGTGCCAAGCTGAGGATCCCA 2425  
Qy 2150 CCTGCAAGATTCAGACCTGATGTTTGAATGCAAGATTACCCCTGTGTGCGACAC 2209  
Db 2426 CCTGCAAGATTCAGACCTGATGTTTGAATGCAAGATTACCCCTGTGTGCGACAC 2485  
Qy 2210 TCTGAGATTCGTATGAGGCAACCGTAACTTAAGATCCCTCACTTGGAGGCAACCA 2269  
Db 2486 TCTGAGATTCGTATGAGGCAACCGTAACTTAAGATCCCTCACTTGGAGGCAACCA 2545  
Qy 2270 TGAAGAAAGAGATGTAAGATGCGGTGTGAAGCTTAAACACCCAAATGTTTGTGG 2329  
Db 2546 TGAAGAAAGAGATGTAAGATGCGGTGTGAAGCTTAAACACCCAAATGTTTGTGG 2605  
Qy 2330 AGTCTTTGAGGCTGAGATTTGCTGTGAGTTGACCCATGCTGTGACTGAAAGATCTCCCAA 2389  
Db 2606 AGTCTTTGAGGCTGAGATTTGCTGTGAGTTGACCCATGCTGTGACTGAAAGATCTCCCAA 2665  
Qy 2390 TCTTTAGACCTCCCGACGCTGAAATCTGTGAGCTGCGAGGAAACAAAGTGAACAGC 2449  
Db 2666 TCTTTAGACCTCCCGACGCTGAAATCTGTGAGCTGCGAGGAAACAAAGTGAACAGC 2725  
Qy 2450 AGGAGATATGCTCTGATGATGCTCTTGAAGTCTCCAGTGGCTCTGCAAGAGCTGA 2509  
Db 2726 AGGAGATATGCTCTGATGATGCTCTTGAAGTCTCCAGTGGCTCTGCAAGAGCTGA 2785  
Qy 2510 TACTGAGAGACTGAGGATCAACGCAAGGTTGCGAGGCTGAGGCTCAAGCCCTCGCA 2569  
Db 2786 TACTGAGAGACTGAGGATCAACGCAAGGTTGCGAGGCTGAGGCTCAAGCCCTCGCA 2845  
Qy 2570 GCAACCGGAGCTTGAACACCTGTGCTATCCAAACAGCCTGGGAAACGAAGTGTAA 2629  
Db 2846 GCAACCGGAGCTTGAACACCTGTGCTATCCAAACAGCCTGGGAAACGAAGTGTAA 2905  
Qy 2630 ATCTACTGTGTGATCCATGAGGCTTCCCACTGTAGTCTGAGAGGCTGATGCTGAATC 2689  
Db 2906 ATCTACTGTGTGATCCATGAGGCTTCCCACTGTAGTCTGAGAGGCTGATGCTGAATC 2965  
Qy 2690 AGTGCCACCTGAGAACAGGCTGAGGCTTCTTGAACCTTGAAGCTTAAGGATTAATCAT 2749  
Db 2966 AGTGCCACCTGAGAACAGGCTGAGGCTTCTTGAACCTTGAAGCTTAAGGATTAATCAT 3025  
Qy 2750 GGTGAGGCACTGAGCTTAACTGATGAACCTGTGGAAGCAATAGCGGTGAAGCTTCTGT 2809  
Db 3026 GGTGAGGCACTGAGCTTAACTGATGAACCTGTGGAAGCAATAGCGGTGAAGCTTCTGT 3085  
Qy 2810 GCGAGGTATGAGAGAACATCTTGTGATCTCCAGAGCTTGAAGTGTAAAGTGTATC 2869  
Db 3086 GCGAGGTATGAGAGAACATCTTGTGATCTCCAGAGCTTGAAGTGTAAAGTGTATC 3145  
Qy 2870 TCACCGCGCGTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGA 2929  
Db 3146 TCACCGCGCGTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACACCTGAAGA 3205





Db 1466 AGAGGTGCCCCCTTCAACCAACGCTCACAGGCTCAGCCCTTTTGCTGATC 1525  
Qy 1250 AGTCACCCCTCGAGCGCGGTCCGGCGCTGCTCAATCTGGAGAAAGATTGTCGA 1309  
Db 1526 AGCTACCCCTCGAGCGGTGTCGGCGCTGCTCAATCTGGAGAAAGATTGCTGGA 1585  
Qy 1310 AGCGCTTGCCGTAAGCTGTGAGGGAATGTGAAATAGAAAGTCAAGTGTGATGAG 1369  
Db 1586 AGCGCTTGCCGTAAGCTGTGAGGGAATGTGAAATAGAAAGTCAAGTGTGATGAG 1645  
Qy 1370 AGGACTCATGTTTAAAGACTCGGGGAGTCTGAGCTCCGTCTGTTTTCATGAACA 1429  
Db 1646 AGGACTCATGTTTAAAGACTCGGGGAGTCTGAGCTCCGTCTGTTTTCATGAACA 1705  
Qy 1430 TCCCTTCCCAAGACAGCCCTGTGAGAGTACTACACCTTCTTCCACCTCAGTCCAGG 1489  
Db 1706 TCTCTTCCCAAGACAGCCCTGTGAGAGTACTACACCTTCTTCCACCTCAGTCCAGG 1765  
Qy 1490 ACTTCTGCGCGCTTGTACTACGTGTAGAGGCGCTGAAATCGAGCAGCTCTGCGC 1549  
Db 1766 ACTTCTGCGCGCTTGTACTACGTGTAGAGGCGCTGAAATCGAGCAGCTCTGCGC 1825  
Qy 1550 CTCTGTACGTTGAGAAACAAAGAGCTCAGTGAAGCTTAAACAGGCGCTTCCATATC 1609  
Db 1826 CTCTGTACGTTGAGAAACAAAGAGCTCAGTGAAGCTTAAACAGGCGCTTCCATATC 1885  
Qy 1610 ACTGCTTTGAGATGAAGGCTTCTGTTTGCGCTGTGAGGGAACCTTAAGAGAGCCAC 1669  
Db 1886 ACTGCTTTGAGATGAAGGCTTCTGTTTGCGCTGTGAGGGAACCTTAAGAGAGCCAC 1945  
Qy 1670 TGGAGGTCCTGTGGGCTGTCCCGTTCCTGTGGGGGTGAGAGAGGCTTGTGACTGGG 1729  
Db 1946 TGGAGGTCCTGTGGGCTGTCCCGTTCCTGTGGGGGTGAGAGAGGCTTGTGACTGGG 2005  
Qy 1730 TCTCTGTGTGGGTGAGAGCTTAAATGCCACCAACCCAGAGACAACCTTGAACGCTTTC 1789  
Db 2006 TCTCTGTGTGGGTGAGAGCTTAAATGCCACCAACCCAGAGACAACCTTGAACGCTTTC 2065  
Qy 1790 ACTGCTTTGAGACTCAAGCAAGAAAGTTGTGCTGTGCACTTAAACAGCTTCCAG 1849  
Db 2066 ACTGCTTTGAGACTCAAGCAAGAAAGTTGTGCTGTGCACTTAAACAGCTTCCAG 2125  
Qy 1850 AAGTGTGCTTCCGATTAACAGAACTGGAACCTTGAATGACATCTTCTTGTGCTCCAGC 1909  
Db 2126 AAGTGTGCTTCCGATTAACAGAACTGGAACCTTGAATGACATCTTCTTGTGCTCCAGC 2185  
Qy 1910 ACTGTCCGTATTTGCGAAATTCGGGTGATGTCAAAAGGATCTTCCCAAGATGAGT 1969  
Db 2186 ACTGTCCGTATTTGCGAAATTCGGGTGATGTCAAAAGGATCTTCCCAAGATGAGT 2245  
Qy 1970 CCGCTGAGGCAATGCTGTGTGCTCTTAATGATGCGGGATTAAGACCTCATATGAGAGC 2029  
Db 2246 CCGCTGAGGCAATGCTGTGTGCTCTTAATGATGCGGGATTAAGACCTCATATGAGAGC 2305  
Qy 2030 AGTGGAGAAATTTTGTGCTCAATGCTTGGCAACCAACCAACCTGTGGAGAGCTGAGCCTGG 2089  
Db 2306 AGTGGAGAAATTTTGTGCTCAATGCTTGGCAACCAACCAACCTGTGGAGAGCTGAGCCTGG 2365  
Qy 2090 GCAGCAGATCTCTGACAGAGCGGACATGAAGACCTGTGTGCAAGCTGAGAGATCCA 2149  
Db 2366 GCAGCAGATCTCTGACAGAGCGGACATGAAGACCTGTGTGCAAGCTGAGAGATCCA 2425  
Qy 2150 CCTGTAAATACAGACCTCTGATGTTTAAAGATGCAAGATTAACCTCTGTGTGAGCAGC 2209  
Db 2426 CCTGTAAATACAGACCTCTGATGTTTAAAGATGCAAGATTAACCTCTGTGTGAGCAGC 2485  
Qy 2210 TCTGAGATGCTCATGAGCAACCGTAAACCTTAAGATCCCTCAACTTGGAGGCAACCCAGC 2269  
Db 2486 TCTGAGATGCTCATGAGCAACCGTAAACCTTAAGATCCCTCAACTTGGAGGCAACCCAGC 2545  
Qy 2270 TGAAGAAAGAGATGAAGATGAGTGGCTGGAACCTTAAACACCCAAATGTTTGTGG 2329

Db 2546 TGAAGAAAGAGATGAAGATGAGCGGTGGAAGCCTTAAACACCCAAATGTTTGTGG 2605  
Qy 2330 AGCTTTTGAAGGCTGATTTGTGCTGATTAAGCCCATGCTGTAACTGAAGATCTCCCAA 2389  
Db 2606 AGCTTTTGAAGGCTGATTTGTGCTGATTAAGCCCATGCTGTAACTGAAGATCTCCCAA 2665  
Qy 2390 TCTTTAAGACCTCCCGAGCCTGAATCTGTGAGCTTGTGGCAGGAAACAGGTGACAGAC 2449  
Db 2666 TCTTTAAGACCTCCCGAGCCTGAATCTGTGAGCTTGTGGCAGGAAACAGGTGACAGAC 2725  
Qy 2450 AGGAGATTAATGCTCTGATGAATGCTTGAAGATCTCCAGTGGCCCTGAGAGAGCTGA 2509  
Db 2726 AGGAGATTAATGCTCTGATGAATGCTTGAAGATCTCCAGTGGCCCTGAGAGAGCTGA 2785  
Qy 2510 TACTGAGAGACTGTGAGATCAACAGCAGGTTTCCAGAGTCTGAGCTCAGCCTGCTCA 2569  
Db 2786 TACTGAGAGACTGTGAGATCAACAGCAGGTTTCCAGAGTCTGAGCTCAGCCTGCTCA 2845  
Qy 2570 GCAACCGGAGCTTGAACAACCTGTGCTTATCAACAAACAGCTGGGGAACGAGGTGTA 2629  
Db 2846 GCAACCGGAGCTTGAACAACCTGTGCTTATCAACAAACAGCTGGGGAACGAGGTGTA 2905  
Qy 2630 ATCTACTGTGTGATCCATGAGGCTTCCCACTGTACTGTGCAAGAGCTGATGCTGAATC 2689  
Db 2906 ATCTACTGTGTGATCCATGAGGCTTCCCACTGTACTGTGCAAGAGCTGATGCTGAATC 2965  
Qy 2690 AGTGCACCTGGAACAGGCTGTGTGTGTTTCTTGAACCTTGACCTTATGGGTAACTCAT 2749  
Db 2966 AGTGCACCTGGAACAGGCTGTGTGTGTTTCTTGAACCTTGACCTTATGGGTAACTCAT 3025  
Qy 2750 GGTGACGACCTGAGCCTTATGACATGAACCTGTGGAAGCAATGAGCTGTGAAGCTTGT 2809  
Db 3026 GGTGACGACCTGAGCCTTATGACATGAACCTGTGGAAGCAATGAGCTGTGAAGCTTGT 3085  
Qy 2810 GCGAGGTCAATGAGAAACCAATCTTGTCAATCTCCAGAGACTGGAATTTGGTAAAGTCAATC 2869  
Db 3086 GCGAGGTCAATGAGAAACCAATCTTGTCAATCTCCAGAGACTGGAATTTGGTAAAGTCAATC 3145  
Qy 2870 TCACCGCGCGGTGTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACCTGAAGA 2929  
Db 3146 TCACCGCGCGGTGTGTGAGAGTCTGTCTGTGTGATCTCGAGGAGCAGACCTGAAGA 3205  
Qy 2930 GCTGTGATCTTACGGAACATGCTGTGGGTGAACGCTGTGGCTGTGCGAGGAGAC 2989  
Db 3206 GCTGTGATCTTACGGAACATGCTGTGGGTGAACGCTGTGGCTGTGCGAGGAGAC 3265  
Qy 2990 TGAAGCAAAAGAACAGTGTGTGAAGAGACTCGGGTTGAAGGCAATGTGACTGACTTCTG 3049  
Db 3266 TGAAGCAAAAGAACAGTGTGTGAAGAGACTCGGGTTGAAGGCAATGTGACTGACTTCTG 3325  
Qy 3050 ATTGCTGTGAGGCACTGCTCCTTGTGGCCCTTCTGTGCAACCGGCACTGTGACGCTTAAAC 3109  
Db 3326 ATTGCTGTGAGGCACTGCTCCTTGTGGCCCTTCTGTGCAACCGGCACTGTGACGCTTAAAC 3385  
Qy 3110 TGTGTGAGATTAATCTTCAAGTCCCAAGAAATGATGAAGCTGTGTGCGCCTTTCCTGTC 3169  
Db 3386 TGTGTGAGATTAATCTTCAAGTCCCAAGAAATGATGAAGCTGTGTGCGCCTTTCCTGTC 3445  
Qy 3170 CCACGTCTAATTAACAGATTAATTTGGGCTGTGGAATTTGCGAGTACCTGTGTGCAATTAAGA 3229  
Db 3446 CCACGTCTAATTAACAGATTAATTTGGGCTGTGGAATTTGCGAGTACCTGTGTGCAATTAAGA 3505  
Qy 3230 AGCTGTGAGGAGTCAAGTCAAGTCAAGCCCGAGTCCGTAATGAGAGTGTGGGATTT 3289  
Db 3506 AGCTGTGAGGAGTCAAGTCAAGTCAAGCCCGAGTCCGTAATGAGAGTGTGGGATTT 3565  
Qy 3290 CTTTGTGATGAAGATGACCG 3308  
Db 3566 CTTTGTGATGAAGATGACCG 3584

RESULT 12  
ABK48609



ID ABK48609 standard; cDNA; 1157 BP.  
XX  
AC ABK48609;  
XX  
DT 13-AUG-2002 (first entry)  
XX  
DE Human MATER cDNA fragment #1.  
XX  
KW Human; gene; 88; contraceptive; antifertility; MATER;  
KW Maternal antigen that embryos require; MATER null phenotype; oocyte;  
KW early embryonic survival; premature ovarian failure; POF;  
KW autoimmune infertility; chromosome 19; gene therapy; fertility.  
XX  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 1..1157  
FT /tag= a  
FT /product= "Human MATER"  
FT /partial  
FT /note= "No start or stop codon shown"  
XX  
PN WO200232955-A1.  
XX  
PD 25-APR-2002.  
XX  
PP 04-APR-2001; 2001WO-US010981.  
XX  
PR 18-OCT-2000; 2000US-0241510P.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Nelson LM, Tong Z;  
XX  
DR WPI; 2002-454595/48.  
XX  
XX P-PsDB; AAU79523.  
XX  
PT New isolated human Maternal Antigen That Embryos Require protein and  
PT polynucleotide, useful in diagnosing or treating fertility or reduced  
PT fertility, or as a contraceptive.  
XX  
PS Claim 11; Page 68-69; 93pp; English.  
XX  
CC The invention discloses an isolated human MATER (Maternal Antigen That  
CC Embryos Require) protein, which can complement a MATER null phenotype in  
CC which zygotes arising from the oocyte do not progress beyond the two-cell  
CC stage. MATER is required for early embryonic survival and abnormal levels  
CC of the protein can lead to premature ovarian failure (POF) and can be  
CC caused by under or over expression of MATER or an autoimmune response to  
CC MATER. MATER is a single-copy Maternal effect gene found on chromosome  
CC 19. The MATER protein and polynucleotide, by gene therapy, are useful in  
CC diagnosing or treating fertility and reduced fertility. In particular,  
CC the MATER protein is useful as a contraceptive agent, or for influencing  
CC (either inhibiting or enhancing) fertility and can be used to detect a  
CC predisposition to infertility or reduced fertility, or for presymptomatic  
CC screening of an individual for infertility/reduced fertility. The protein  
CC and polynucleotide are also useful for detecting an excess or deficiency,  
CC or genetic mutation, of the MATER protein in a mammalian subject (e.g. a  
CC human or a mouse) or for screening for a compound useful in influencing  
CC MATER-mediated fertility. The sequence presented is the human MATER cDNA  
CC fragment #1, which was isolated from a human cDNA library  
XX  
SQ Sequence 1157 BP; 240 A; 324 C; 320 G; 273 T; 0 U; 0 Other;  
Query Match 27.3%; Score 1102; DB 6; Length 1157;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 757 GCTCCGTCGACGAGATCATGTCCTCCGACGAGAGGCTGTTTATCATATTGACGGTTTC 816  
DB 4 GCTCCGTCGACGAGATCATGTCCTCCGACGAGAGGCTGTTTATCATATTGACGGTTTC 63  
QY 817 GATGACCTGGGCTGTCTCTCAATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAG 876

DB 64 GATGACCTGGGCTGTCTCTCAATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAG 123  
QY 877 CAGCTCCGTCACCTCATACGAGTGTCTGAGAGAGCTCTGCTCTGAGTCTTC 936  
DB 124 CAGCTCCGTCACCTCATACGAGTGTCTGAGAGAGCTCTGCTCTGAGTCTTC 183  
QY 937 CTGATCGTACCGTACAGAGAGCTGGGACACAGAGAGCTCAATCAGAGTCTGTCTCC 996  
DB 184 CTGATCGTACCGTACAGAGAGCTGGGACACAGAGAGCTCAATCAGAGTCTGTCTCC 243  
QY 997 CATTACCTGTGTAGAGAGATCTCCGGGAGCAAGAGATCACTTGCTCTTGAGGCG 1056  
DB 244 CATTACCTGTGTAGAGAGATCTCCGGGAGCAAGAGATCACTTGCTCTTGAGGCG 303  
QY 1057 GGGATTTGGTGAACATCAGAGAGACAGAGAGGTTGCTGCGATCATGAACCGTGAAGCTG 1116  
DB 304 GGGATTTGGTGAACATCAGAGAGACAGAGAGGTTGCTGCGATCATGAACCGTGAAGCTG 363  
QY 1117 CTGACACAGTGCAGAGTCCCGGCGCTCTCATCTGCTGGTCCCTGAGCTGCAG 1176  
DB 364 CTGACACAGTGCAGAGTCCCGGCGCTCTCATCTGCTGGTCCCTGAGCTGCAG 423  
QY 1177 GACGTGTGGGGAGAGAGCGTCCGCTTCAACCAAGGCTCAAGAGGCTGAGCGCGCT 1236  
DB 424 GACGTGTGGGGAGAGAGCGTCCGCTTCAACCAAGGCTCAAGAGGCTGAGCGCGCT 483  
QY 1237 TTTGTGTTTCATCAGCTCACCCCTTGAGAGCGTGTCCGGCGCTGTCTCATCTGAGGAA 1296  
DB 484 TTTGTGTTTCATCAGCTCACCCCTTGAGAGCGTGTCCGGCGCTGTCTCATCTGAGGAA 543  
QY 1297 AGAGTTGTCTGAGAGCGCTTCTGCGGTATGCTGTGAGAGAGTGTGAATGGAATCA 1356  
DB 544 AGAGTTGTCTGAGAGCGCTTCTGCGGTATGCTGTGAGAGAGTGTGAATGGAATCA 603  
QY 1357 GGTGTTGATGGAGAGAGCTCATGTTCAAGAGAGCTGGGAGTCAAGCTCCGCTCTG 1416  
DB 604 GGTGTTGATGGAGAGAGCTCATGTTCAAGAGAGCTGGGAGTCAAGCTCCGCTCTG 663  
QY 1417 TTTCAATGAACATCTCTCTCCAGACAGCACTGTGAGAGTACTACCTCTTCCAC 1476  
DB 664 TTTCAATGAACATCTCTCTCCAGACAGCACTGTGAGAGTACTACCTCTTCCAC 723  
QY 1477 CTGAGTCTCAGAGATCTCTGTGCGGCTTGTACTAGTGTGAGAGGCTGAGAAATCGAG 1536  
DB 724 CTGAGTCTCAGAGATCTCTGTGCGGCTTGTACTAGTGTGAGAGGCTGAGAAATCGAG 783  
QY 1537 CGAGTCTCTGCGCTCTGTAGTGAAGAGCAAGAGAGTCCATGAGACTTAAACAGGCA 1596  
DB 784 CGAGTCTCTGCGCTCTGTAGTGAAGAGCAAGAGAGTCCATGAGACTTAAACAGGCA 843  
QY 1597 GGCCTTCATATCACTGCTTGTGATGAAGAGTCTTGTGTGGCTGTGAGCGAAGAC 1656  
DB 844 GGCCTTCATATCACTGCTTGTGATGAAGAGTCTTGTGTGGCTGTGAGCGAAGAC 903  
QY 1657 GTAAGAGAGCACTGAGAGTCTGTGAGGCTGTGCGGCTTCCCTGAGGAGTGAAGCAAG 1716  
DB 904 GTAAGAGAGCACTGAGAGTCTGTGAGGCTGTGCGGCTTCCCTGAGGAGTGAAGCAAG 963  
QY 1717 CTTCGCACTGGGCTCTCTGTGGGTGAAGAGCTTAATGCAACACCCAGAGAGACCC 1776  
DB 964 CTTCGCACTGGGCTCTCTGTGGGTGAAGAGCTTAATGCAACACCCAGAGAGACCC 1023  
QY 1777 CTGAGAGGCTTCACTGCTTGTGAGAGCTCAAGCAAGAGTGTGCTGCTGGCATTTA 1836  
DB 1024 CTGAGAGGCTTCACTGCTTGTGAGAGCTCAAGCAAGAGTGTGCTGCTGGCATTTA 1083  
QY 1837 AACAGCTTCAAGAGAGTGTGCTTCCATTAACCAAGAACTGAGCTTGATGATCTTCC 1896  
DB 1084 AACAGCTTCAAGAGAGTGTGCTTCCATTAACCAAGAACTGAGCTTGATGATCTTCC 1143  
QY 1897 TTTGCTCTCCAGC 1909

Db	1144	TTCTGCTCCAGC	1156
RESULT 13			
AA048999			
ID	AA048999	standard; cDNA; 1157 BP.	
XX			
AC	AA048999;		
DT	07-MAR-2003	(first entry)	
XX			
DE	Human MATER cDNA fragment 1.		
XX			
KM	Human; MATER protein; infertility; fertility; contraceptive agent;		
XX	gene therapy; gene; ss.		
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	1..1155		
FT	/*tag= a		
FT	/product= "Human MATER protein fragment 1"		
FT	/transl_except= (pos:112..114, aa:Thr)		
FT	/transl_except= (pos:589..591, aa:Thr)		
FT	/transl_except= (pos:865..867, aa:Thr)		
FT	/transl_except= (pos:973..975, aa:Thr)		
FT	/transl_except= (pos:1102..1104, aa:Thr)		
FT	/note= "No start and stop codon"		
FT	/partial		
XX			
PN	WO200281492-A1.		
XX			
PD	17-OCT-2002.		
XX			
PF	29-MAR-2002; 2002WO-US009776.		
XX			
PR	04-APR-2001; 2001WO-US010981.		
XX			
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICES.		
PI	Nelson LM, Tong Z;		
XX			
PI	WPI: 2003-058494/05.		
DR	P-ESDB; AME31746.		
XX			
XX	New isolated variant MATER proteins and nucleic acids, useful for		
PT	diagnosing, prognosing and treating infertility and reduced fertility,		
PT	and as contraceptive agents.		
XX			
PS	Example 2; Page 79-80; 110pp; English.		
XX			
CC	The present invention relates to novel MATER proteins and polynucleotides		
CC	encoding such proteins. The MATER proteins are essential to fertility.		
CC	Sequences of the invention are useful for diagnosing, prognosing and		
CC	treating infertility, reduced fertility and as contraceptive agents. They		
CC	are also useful in gene therapy. The method is useful for detecting a		
CC	preposition to or pre-symptomatic screening of an individual for		
CC	infertility or reduced fertility. The present sequence is human MATER		
CC	cDNA fragment 1		
XX			
SO	Sequence 1157 BP; 240 A; 324 C; 320 G; 273 T; 0 U; 0 Other;		
XX			
Query Match	27.3%; Score 1102; DB 8; Length 1157;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 1152;	Conservative 0; Mismatches 1; Indels 0; Gaps 0		
0Y	757 GCTCGGTGACGAGATCATGTCCGACCAAGAGGCTGTTCATCATTTGACGGTTTC	816	
Db	4 GCTCGGTGACGAGATCATGTCCGACCAAGAGGCTGTTCATCATTTGACGGTTTC	63	
0Y	817 GATGACCTGGGCTGTCTCTCAACATGACACAAAGCTTGTGAAAGATCGGCTGAGAG	876	
Db	64 GATGACCTGGGCTGTCTCTCAACATGACACAAAGCTTGTGAAAGATCGGCTGAGAG	123	

QY	877	CAGCCTCCGTTCAACCTTCATATGCGAGTCTGTGGAGAAAGTCCGTCCCGTAGGCTCTTC	936
Db	124	CAGCCTCCGTTCAACCTTCATATGCGAGTCTGTGGAGAAAGTCCGTCCCGTAGGCTCTTC	183
QY	937	CTGATTCGCAACCGTCAGAGAGCTGTGGGCAACAGAAAGCTCAAGTCTAGAGGTGTGTCTCC	996
Db	184	CTGATTCGCAACCGTCAGAGAGCTGTGGGCAACAGAAAGCTCAAGTCTAGAGGTGTGTCTCC	243
QY	997	CGTTAACCTGTTAGTTAGAGAAATCTCCGGGGAAACAAGAAATCCAATTGCTCTTTGAGCGC	1056
Db	244	CGTTAACCTGTTAGTTAGAGAAATCTCCGGGGAAACAAGAAATCCAATTGCTCTTTGAGCGC	303
QY	1057	GGGATTTGGTAGCATCAAAACACACAAGAGTTGGGTGGATCATGAAACAACTGTGAGCTG	1116
Db	304	GGGATTTGGTAGCATCAAAACACACAAGAGTTGGGTGGATCATGAAACAACTGTGAGCTG	363
QY	1117	CTGCACCAAGTACCAGAGTCCCGCGTGGGCTCTCTCATCTGCGTGGCCCTGACGTGACG	1176
Db	364	CTGCACCAAGTACCAGAGTCCCGCGTGGGCTCTCTCATCTGCGTGGCCCTGACGTGACG	423
QY	1177	GACGTGTGGGGGAGAGAGCTGCACCCTTCAACAAACGCTCACAGGCTTGACAGCGCGCT	1236
Db	424	GACGTGTGGGGGAGAGAGCTGCACCCTTCAACAAACGCTCACAGGCTTGACAGCGCGCT	483
QY	1237	TTTGTGTTTCATCAAGCTCAACCCCTCGAGAGCGTGGTCCGGCGCTGTCTCAATCTGGAGAA	1296
Db	484	TTTGTGTTTCATCAAGCTCAACCCCTCGAGAGCGTGGTCCGGCGCTGTCTCAATCTGGAGAA	543
QY	1297	AGAGTGTGCTGAAGCGCTTCTGCGATATGGCTGGAGGGAGTGGAAATAGGAAGTCA	1356
Db	544	AGAGTGTGCTGAAGCGCTTCTGCGATATGGCTGGAGGGAGTGGAAATAGGAAGTCA	603
QY	1357	GTCGTTTGAAGTGTGACGACCTCATGAGTTCAAGAACTCGGGAGTGTGAGCTCCGTGCTTG	1416
Db	604	GTCGTTTGAAGTGTGACGACCTCATGAGTTCAAGAACTCGGGAGTGTGAGCTCCGTGCTTG	663
QY	1417	TTTTCACATGAACATCTCTTCTTCCACAGACCACTGTGAGAGTACTACACTTCTTCCAC	1476
Db	664	TTTTCACATGAACATCTCTTCTTCCACAGACCACTGTGAGAGTACTACACTTCTTCCAC	723
QY	1477	CTCAGTCTCCAGAGCTTCTGTGCGGCTTGTATCTACGTTTAGAGGGCTGGAAATCGAG	1536
Db	724	CTCAGTCTCCAGAGCTTCTGTGCGGCTTGTATCTACGTTTAGAGGGCTGGAAATCGAG	783
QY	1537	CCAGCTCTGACCTCTGTAGTGTGAGAGACAAGAGAGTCCATGAGACTTAAACAGCA	1596
Db	784	CCAGCTCTGACCTCTGTAGTGTGAGAGACAAGAGAGTCCATGAGACTTAAACAGCA	843
QY	1597	GAGTTCCATATCACTCGCTTTGTGATGAAGCGTTCTTTGTTGGCTCTGTGAGCGAAGAC	1656
Db	844	GAGTTCCATATCACTCGCTTTGTGATGAAGCGTTCTTTGTTGGCTCTGTGAGCGAAGAC	903
QY	1657	GTAAGGAGGCCACTGTGAGAGTCTGTGTGGGCTGTCCCTTCCCTGGGGGTGAACGACGAG	1716
Db	904	GTAAGGAGGCCACTGTGAGAGTCTGTGTGGGCTGTCCCTTCCCTGGGGGTGAACGACGAG	963
QY	1717	CTTCGCACTGGGTCTCTGTGTGGGTACGAGCGCTAATGCCACAACCCAGAGACACC	1776
Db	964	CTTCGCACTGGGTCTCTGTGTGGGTACGAGCGCTAATGCCACAACCCAGAGACACC	1022
QY	1777	CTGAGCGCTTCACTGTCTTTTGGAGACTCAAGACAAGATTTGTTGCTTG6CATTA	1836
Db	1024	CTGAGCGCTTCACTGTCTTTTGGAGACTCAABACAAAGATTTGTTGCGTTTG6CATTA	1088
QY	1837	AACAGCTTCCAAGAAGTGTGGCTTCCGATTAAACAAGAACTTGACCTTGATATGCAATCTTCC	1896
Db	1084	AACAGCTTCCAAGAAGTGTGGCTTCCGATTAAACAAGAACTTGACCTTGATATGCAATCTTCC	1144
QY	1897	TTTCGTCTCCAGC 1909	
Db	1144	TTTCGTCTCCAGC 1156	





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2005, 01:48:06 ; Search time 633 Seconds

(without alignments)  
10430.295 Million cell updates/sec

Title: US-10-066-521-5

Perfect score: 4035

Sequence: 1 atggaagagagaacaatcgct.....tcgaggcgctgctcttaaa 4035

Scoring table: OLIGO\_NUC

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTCUTS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
C 1	20	0.5	601	US-09-949-016-91585	Sequence 91585, A
C 2	20	0.5	601	US-09-949-016-91586	Sequence 91586, A
C 3	20	0.5	6042	US-08-261-822A-1	Sequence 1, Appl
C 4	20	0.5	6042	PCT-US95-07744A-1	Sequence 1, Appl
C 5	20	0.5	6172	US-08-819-288-1	Sequence 1, Appl
C 6	20	0.5	6172	US-09-400-348-1	Sequence 1, Appl
C 7	20	0.5	8733	US-10-144-198-11	Sequence 11, Appl
C 8	20	0.5	57837	US-09-949-016-13371	Sequence 14371, A
C 9	20	0.5	57839	US-09-949-016-17601	Sequence 17601, A
C 10	19	0.5	400	US-09-621-976-2835	Sequence 2835, Ap
C 11	19	0.5	601	US-09-949-016-19079	Sequence 19079, A
C 12	19	0.5	601	US-09-949-016-19080	Sequence 19080, A
C 13	19	0.5	601	US-09-949-016-36893	Sequence 36893, A
C 14	19	0.5	601	US-09-949-016-36894	Sequence 36894, A
C 15	19	0.5	601	US-09-949-016-58788	Sequence 58788, A
C 16	19	0.5	601	US-09-949-016-151197	Sequence 151197, A
C 17	19	0.5	601	US-09-949-016-151198	Sequence 151198, A
C 18	19	0.5	601	US-09-949-016-189239	Sequence 189239, A
C 19	19	0.5	601	US-09-949-016-189240	Sequence 189240, A
C 20	19	0.5	601	US-09-949-016-189279	Sequence 189279, A
C 21	19	0.5	601	US-09-949-016-189280	Sequence 189280, A
C 22	19	0.5	1110	US-09-543-681A-2553	Sequence 2553, Ap
C 23	19	0.5	1341	US-09-513-999C-14927	Sequence 14927, A
C 24	19	0.5	1428	US-09-248-796A-3252	Sequence 3252, Ap
C 25	19	0.5	2848	US-09-023-655-1364	Sequence 1364, Ap
C 26	19	0.5	3180	US-09-248-796A-20	Sequence 20, Appl
C 27	19	0.5	3349	US-09-375-318-36	Sequence 36, Appl

## ALIGNMENTS

28	19	0.5	5521	US-08-956-171E-408	Sequence 408, App
29	19	0.5	5521	US-08-781-986A-408	Sequence 408, App
30	19	0.5	13198	US-09-949-016-16425	Sequence 16425, A
31	19	0.5	28843	US-09-949-016-11812	Sequence 11812, A
32	19	0.5	28843	US-09-949-016-12712	Sequence 12712, A
33	19	0.5	28843	US-09-949-016-17158	Sequence 17158, A
34	19	0.5	28843	US-09-949-016-17158	Sequence 17159, A
35	19	0.5	48039	US-09-949-016-15990	Sequence 15990, A
36	19	0.5	49315	US-09-949-016-13016	Sequence 13016, A
37	19	0.5	56326	US-09-949-016-16468	Sequence 16468, A
38	19	0.5	57103	US-09-949-016-13445	Sequence 13445, A
39	19	0.5	60417	US-09-949-016-13312	Sequence 13312, A
40	19	0.5	71251	US-09-949-016-15332	Sequence 15332, A
41	19	0.5	209210	US-09-949-016-15094	Sequence 15094, A
42	19	0.5	278866	US-09-949-016-13922	Sequence 13922, A
43	19	0.5	278866	US-09-949-016-13923	Sequence 13923, A
44	19	0.5	278866	US-09-949-016-13924	Sequence 13924, A
45	19	0.5	278866	US-09-949-016-13925	Sequence 13925, A

```

RESULT 1
US-09-949-016-91585/c
; Sequence 91585, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 91585
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-91585

Query Match      0.5%; Score 20; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2398 ACCCTCCCGAGCGTGAATC 2417
Db      476 ACCCTCCCGAGCGTGAATC 457

RESULT 2
US-09-949-016-91586/c
; Sequence 91586, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```

PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 91586  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-91586

Query Match 0.5%; Score 20; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2398 ACCTCCCCAGCTGAATC 2417  
DB 290 ACCTCCCCAGCTGAATC 271

## RESULT 3

US-08-261-822A-1/c  
Sequence 1, Application US/08261822A  
Patent No. 5650553

GENERAL INFORMATION:  
APPLICANT: Becker, Joseph R. et al.  
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene  
TITLE OF INVENTION: and Pathogens  
NUMBER OF SEQUENCES: 82  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553r1s  
STREET: One Liberty Place, 46th floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/261,822A  
FILING DATE: 17-JUN-1994  
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6042 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-261-822A-1

Query Match 0.5%; Score 20; DB 1; Length 6042;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 966 AGAGAGCTCAAGTCAGAG 985  
DB 2046 AGAGAGCTCAAGTCAGAG 2027

## RESULT 4

PCT-US95-07744A-1/c  
Sequence 1, Application PC/TUS9507744A  
GENERAL INFORMATION:

APPLICANT: Trustees of The University of Pennsylvania  
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene  
TITLE OF INVENTION: and Pathogens  
NUMBER OF SEQUENCES: 82  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris  
STREET: One Liberty Place, 46th floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07744A  
FILING DATE: 15-JUNE-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/261,822  
FILING DATE: June 17, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6042 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
PCT-US95-07744A-1

Query Match 0.5%; Score 20; DB 5; Length 6042;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 966 AGAGAGCTCAAGTCAGAG 985  
DB 2046 AGAGAGCTCAAGTCAGAG 2027

RESULT 5  
US-08-819-288-1/c  
Sequence 1, Application US/08819288  
Patent No. 5955652

GENERAL INFORMATION:  
APPLICANT: Becker, Joseph  
APPLICANT: Alonso, Jose  
TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE  
TITLE OF INVENTION: AND PATHOGENS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5955652r1s  
STREET: One Liberty Place - 46th floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/819,288

FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
REFERENCE/DOCKET NUMBER: UPN-2949  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6172 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-819-288-1

Query Match 0.5%; Score 20; DB 2; Length 6172;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 966 AGAGAGCTCAAGTCAGAG 985  
Db 1975 AGAGAGCTCAAGTCAGAG 1956

RESULT 6  
US-09-400-348-1/c  
Sequence 1, Application US/09400348  
Patent No. 6355778  
GENERAL INFORMATION:  
APPLICANT: Ecker, Joseph  
APPLICANT: Alonso, Jose  
TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE  
TITLE OF INVENTION: AND PATHOGENS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6355778ris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/400,348  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/819,288  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
REFERENCE/DOCKET NUMBER: UPN-2949  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6172 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO

ANTI-SENSE: NO  
US-09-400-348-1

Query Match 0.5%; Score 20; DB 3; Length 6172;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 966 AGAGAGCTCAAGTCAGAG 985  
Db 1975 AGAGAGCTCAAGTCAGAG 1956

RESULT 7  
US-10-144-198-11  
Sequence 11, Application US/10144198  
Patent No. 6833247  
GENERAL INFORMATION:  
APPLICANT: Origene Technologies Inc  
TITLE OF INVENTION: Regulated Prostate Cance Genes  
FILE REFERENCE: 9U 105 R1  
CURRENT APPLICATION NUMBER: US/10/144,198  
CURRENT FILING DATE: 2002-05-14  
NUMBER OF SEQ ID NOS: 44  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 11  
LENGTH: 8733  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (473)..(8566)  
US-10-144-198-11

Query Match 0.5%; Score 20; DB 4; Length 8733;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 322 ACAGCAGCAGCAGCAGAGA 341  
Db 8459 ACAGCAGCAGCAGCAGAGA 8478

RESULT 8  
US-09-949-016-14371/c  
Sequence 14371, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CLO01307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14371  
LENGTH: 57837  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-14371

Query Match 0.5%; Score 20; DB 4; Length 57837;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2398 ACCTCCCCGCGCTGAATC 2417  
|||||



Db 19969 ACCTCCCGCAGCTGAATC 19950

## RESULT 9

US-09-949-016-17601/c  
; Sequence 17601, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17601  
; LENGTH: 57839  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17601

Query Match 0.5%; Score 20; DB 4; Length 57839;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2398 ACCTCCCGCAGCTGAATC 2417  
Db 19970 ACCTCCCGCAGCTGAATC 19951

## RESULT 10

US-09-621-976-2835  
; Sequence 2835, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 2835  
; LENGTH: 400  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 99..308  
; NAME/KEY: misc\_feature  
; LOCATION: 21  
; OTHER INFORMATION: n=a, g, c or t  
US-09-621-976-2835

Query Match 0.5%; Score 19; DB 4; Length 400;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3573 AGCCGTGCTGAAGTCT 3591  
Db 333 AGCCGTGCTGAAGTCT 351

## RESULT 11

US-09-949-016-19079  
; Sequence 19079, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 19079  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-19079

Query Match 0.5%; Score 19; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1468 TTCTTCACCTCAGCTTCC 1486  
Db 29 TTCTTCACCTCAGCTTCC 47

RESULT 12  
US-09-949-016-19080  
; Sequence 19080, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 19080  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-19080

Query Match 0.5%; Score 19; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1468 TTCTTCACCTCAGCTTCC 1486  
Db 41 TTCTTCACCTCAGCTTCC 59

RESULT 13  
US-09-949-016-36893  
; Sequence 36893, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36893
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-36893

Query Match          0.5%: Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1468 TTCTTCACCTCAGTCTCC 1486
Db      29  TTCTTCACCTCAGTCTCC 47

RESULT 14
US-09-949-016-36894
; Sequence 36894, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36894
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-36894

Query Match          0.5%: Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1468 TTCTTCACCTCAGTCTCC 1486
Db      41  TTCTTCACCTCAGTCTCC 59

RESULT 15
US-09-949-016-58788
; Sequence 58788, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58788
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-58788

Query Match          0.5%: Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2416 TCTTGAGCTGGCAGGAA 2434
Db      147  TCTTGAGCTGGCAGGAA 165
```

Search completed: July 19, 2005, 11:19:54  
Job time : 637 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2005, 02:27:16 ; Search time 2592 Seconds

(without alignments)  
9876.191 Million cell updates/sec

Title: US-10-066-521-5

Perfect score: 4035

Sequence: 1 atgagagagagaacatcgct.....tcgaggggctgtgtctttaa 4035

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 7173243 seqs, 3172129809 residues

Word size : 0

Total number of hits satisfying chosen parameters: 14346486

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US05\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US04\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US03\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US02\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US01\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US00\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US05\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US04\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US03\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US02\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US01\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US00\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
23: /cgn2\_6/ptodata/1/pubpna/US05\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/1/pubpna/US04\_PUBCOMB.seq.\*  
25: /cgn2\_6/ptodata/1/pubpna/US03\_PUBCOMB.seq.\*  
26: /cgn2\_6/ptodata/1/pubpna/US02\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4035	100.0	4035	US-10-124-498-5	Sequence 5, Appli
2	4035	100.0	4035	US-10-066-521-5	Sequence 5, Appli
3	2917	72.3	3489	US-10-416-642-3	Sequence 3, Appli
4	2917	72.3	3926	US-10-216-645-1	Sequence 1, Appli
5	2862	70.9	3830	US-10-216-645-3	Sequence 3, Appli
6	2801	69.4	3226	US-10-092-900A-347	Sequence 347, App
7	2713	67.2	3885	US-10-860-761-3	Sequence 3, Appli

8	2713	67.2	3900	US-10-399-443-23	Sequence 23, Appli
9	2713	67.2	3900	US-10-677-943-23	Sequence 23, Appli
10	1102	27.3	1157	US-10-399-443-1	Sequence 1, Appli
11	1102	27.3	1157	US-10-677-943-1	Sequence 1, Appli
12	657	16.3	1075	US-10-399-443-3	Sequence 3, Appli
13	657	16.3	1075	US-10-677-943-3	Sequence 3, Appli
14	174	4.3	2099	US-10-027-632-258159	Sequence 258159,
15	174	4.3	2099	US-10-027-632-258159	Sequence 258159,
16	24	0.6	24	US-10-399-443-13	Sequence 13, Appli
17	24	0.6	24	US-10-399-443-13	Sequence 13, Appli
18	24	0.6	24	US-10-092-900A-750	Sequence 750, App
19	24	0.6	24	US-10-677-943-13	Sequence 13, Appli
20	24	0.6	24	US-10-677-943-13	Sequence 13, Appli
21	22	0.5	22	US-10-399-443-7	Sequence 7, Appli
22	22	0.5	22	US-10-677-943-7	Sequence 7, Appli
23	22	0.5	22	US-10-399-443-9	Sequence 9, Appli
24	22	0.5	26	US-10-677-943-9	Sequence 9, Appli
25	22	0.5	27	US-10-399-443-17	Sequence 17, Appli
26	22	0.5	27	US-10-677-943-17	Sequence 17, Appli
27	22	0.5	1036	US-10-739-930-4220	Sequence 4220, Ap
28	22	0.5	3447	US-10-399-443-5	Sequence 5, Appli
29	22	0.5	3447	US-10-677-943-5	Sequence 5, Appli
30	21	0.5	21	US-10-216-645-6	Sequence 6, Appli
31	21	0.5	21	US-10-216-645-10	Sequence 10, Appli
32	21	0.5	21	US-10-399-443-8	Sequence 8, Appli
33	21	0.5	21	US-10-399-443-19	Sequence 19, Appli
34	21	0.5	21	US-10-399-443-20	Sequence 20, Appli
35	21	0.5	21	US-10-092-900A-749	Sequence 749, App
36	21	0.5	21	US-10-092-900A-751	Sequence 751, App
37	21	0.5	21	US-10-677-943-8	Sequence 8, Appli
38	21	0.5	21	US-10-677-943-19	Sequence 19, Appli
39	21	0.5	21	US-10-677-943-20	Sequence 20, Appli
40	21	0.5	21	US-10-677-943-16	Sequence 16, Appli
41	21	0.5	22	US-10-399-443-16	Sequence 16, Appli
42	21	0.5	22	US-10-677-943-16	Sequence 16, Appli
43	21	0.5	476	US-09-918-995-6234	Sequence 6234, Ap
44	21	0.5	483	US-09-918-995-20036	Sequence 20036, A
45	21	0.5	495	US-10-723-860-2423	Sequence 2423, Ap

## ALIGNMENTS

RESULT 1

US-10-124-498-5

Sequence 5, Application US/10124498

Publication No. US20030017983A1

GENERAL INFORMATION:

APPLICANT: Bertin, John

APPLICANT: Wang, Weiye

APPLICANT: Blatcher, Maria

TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR

TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF

FILE REFERENCE: 07334-367001

CURRENT APPLICATION NUMBER: US/10/124,498

CURRENT FILING DATE: 2002-04-17

PRIOR APPLICATION NUMBER: 10/066,521

PRIOR FILING DATE: 2002-01-31

PRIOR APPLICATION NUMBER: 60/318,645

PRIOR FILING DATE: 2001-09-10

PRIOR APPLICATION NUMBER: 60/265,231

NUMBER OF SEQ ID NOS: 25

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5

LENGTH: 4035

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(4032)

US-10-124-498-5

Query Match 100.0%; Score 4035; DB 14; Length 4035;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAGAGCAAAATGCTCACTTTTTCAGCTACGGGCTGCAATGGTGTCTATGAG 60  
DB 1 ATGGAAGAGCAAAATGCTCACTTTTTCAGCTACGGGCTGCAATGGTGTCTATGAG 60

QY 61 CTAGCAAGGAAGATTTTCAACATTCAGAGAAATTAAGAAAGAAATTTTCAAGATC 120  
DB 61 CTAGCAAGGAAGATTTTCAACATTCAGAGAAATTAAGAAAGAAATTTTCAAGATC 120

QY 121 ACCACATGCTCTATTCACAGATTGGAATTCAGAAATGCAAGTGAATGTCTGGCACTC 180  
DB 121 ACCACATGCTCTATTCACAGATTGGAATTCAGAAATGCAAGTGAATGTCTGGCACTC 180

QY 181 CTCTTGCAATGATATTAATGAGCATGCTGGCTTGGGCTACGTCCATTAGCATCTTTGAA 240  
DB 181 CTCTTGCAATGATATTAATGAGCATGCTGGCTTGGGCTACGTCCATTAGCATCTTTGAA 240

QY 241 AACATGAACCTGCGCAACCTCTCGGAGAAAGCAAGGATGACATGAAAAAATTTTCAAA 300  
DB 241 AACATGAACCTGCGCAACCTCTCGGAGAAAGCAAGGATGACATGAAAAAATTTTCAAA 300

QY 301 GCTATGGAACAAGAGGTGCAACAGCAGACAGAGACAGAAATTTTCAAAAGCT 360  
DB 301 GCTATGGAACAAGAGGTGCAACAGCAGACAGAGACAGAAATTTTCAAAAGCT 360

QY 361 ATGGAACAAGAGGTGCAACAGCAGCAGAGACAGAAACAAGACATGAGAGTGAACA 420  
DB 361 ATGGAACAAGAGGTGCAACAGCAGCAGAGACAGAAACAAGACATGAGAGTGAACA 420

QY 421 TGGGACTCAAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGATGACGTGATGT 480  
DB 421 TGGGACTCAAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGATGACGTGATGT 480

QY 481 TTTGAAAAACATGCTGCTGACCTGGCCGGAATGCAAAAGTGGCTGTCTTTGATTC 540  
DB 481 TTTGAAAAACATGCTGCTGACCTGGCCGGAATGCAAAAGTGGCTGTCTTTGATTC 540

QY 541 GACCGGTGGGCTTCGGGCTTCGCAAGGTGTTCTGCAAGGAAAGTCAAGAAATTTGGAAA 600  
DB 541 GACCGGTGGGCTTCGGGCTTCGCAAGGTGTTCTGCAAGGAAAGTCAAGAAATTTGGAAA 600

QY 601 TCGGCTCAAGCAGAAAGATGCTGTGTGCTGGGCGGCAAGTGAATCTTACAGGGAATG 660  
DB 601 TCGGCTCAAGCAGAAAGATGCTGTGTGCTGGGCGGCAAGTGAATCTTACAGGGAATG 660

QY 661 TTTCTCTACGTCTTCTCTCCCGTTAGAGAGATGACGGGATGAGAGAGAGACAGTGT 720  
DB 661 TTTCTCTACGTCTTCTCTCCCGTTAGAGAGATGACGGGATGAGAGAGAGACAGTGT 720

QY 721 AACAAGTTCATTCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
DB 721 AACAAGTTCATTCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

QY 781 GCAACAGAAAGGCTGTGTTATCAATGACGGTTCATGACCTGGGCTGTGCTCTCAAC 840  
DB 781 GCAACAGAAAGGCTGTGTTATCAATGACGGTTCATGACCTGGGCTGTGCTCTCAAC 840

QY 841 AATGACACAAGAGCTTCGAAAGACTGGGCTGAGAGACGCTCCGTTCACTCAAGC 900  
DB 841 AATGACACAAGAGCTTCGAAAGACTGGGCTGAGAGACGCTCCGTTCACTCAAGC 900

QY 901 AGTCTGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
DB 901 AGTCTGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

QY 961 GGCACAGAGAGAGTCAAGTCAAGAGTGTGTCTCCCGCTTACCTGTGTTAGAGGAATC 1020  
DB 961 GGCACAGAGAGAGTCAAGTCAAGAGTGTGTCTCCCGCTTACCTGTGTTAGAGGAATC 1020

QY 1021 TCCGGGAAACAAGAAATCACTGTGCTCTTGAAGCGGGAATTTGAGCATCAAGACA 1080

DB 1021 TCCGGGAAACAAGAAATCACTGTGCTCTTGAAGCGGGAATTTGAGCATCAAGACA 1080

QY 1081 CAAGGTTGCTGCTGATCAAGAACACCGTGAAGTCTGCTGCAACAGTGCAGGTGCCCC 1140  
DB 1081 CAAGGTTGCTGCTGATCAAGAACACCGTGAAGTCTGCTGCAACAGTGCAGGTGCCCC 1140

QY 1141 GTGGGCTCTCATCTGAGTGGCCCTGAGCTGAGAGACGAGTGGGAGAGAGTGGCC 1200  
DB 1141 GTGGGCTCTCATCTGAGTGGCCCTGAGCTGAGAGACGAGTGGGAGAGAGTGGCC 1200

QY 1201 CCTTTCAACAAGAGCTCAAGGCTGCAAGCCCTTTTGTGTTTCAATCAAGCTCAACCCT 1260  
DB 1201 CCTTTCAACAAGAGCTCAAGGCTGCAAGCCCTTTTGTGTTTCAATCAAGCTCAACCCT 1260

QY 1261 CGAGGCTGTGCTGGGCTGCTCTCAATCTGAGAGAAAGTGTGCTGAAAGCCTTCTGC 1320  
DB 1261 CGAGGCTGTGCTGGGCTGCTCTCAATCTGAGAGAAAGTGTGCTGAAAGCCTTCTGC 1320

QY 1321 CGTATGCTGTGAGAGAGTGGGAATAGGAAGTCAAGTGTGATGAGTGAAGCACTCATG 1380  
DB 1321 CGTATGCTGTGAGAGAGTGGGAATAGGAAGTCAAGTGTGATGAGTGAAGCACTCATG 1380

QY 1381 GTTCAAGAGACTCGGAGAGTCTGAGCTCGTCTGTTTCAATGAACATCTCTTCTCCA 1440  
DB 1381 GTTCAAGAGACTCGGAGAGTCTGAGCTCGTCTGTTTCAATGAACATCTCTTCTCCA 1440

QY 1441 GACAGCCACTGTGAGAGTACTACACTTTTCCACTCAAGTCTCCAGACTTGTGACC 1500  
DB 1441 GACAGCCACTGTGAGAGTACTACACTTTTCCACTCAAGTCTCCAGACTTGTGACC 1500

QY 1501 GCTTTGACTACGATTAAGAGGCTGGAATCGAGGAGCTCTGCTGCTGCTGCTGCTGCT 1560  
DB 1501 GCTTTGACTACGATTAAGAGGCTGGAATCGAGGAGCTCTGCTGCTGCTGCTGCTGCT 1560

QY 1561 GAGAAACAAGAGTCAATGAGCTTAAACAGGCAAGCTTCAATATCACTGCTTTGG 1620  
DB 1561 GAGAAACAAGAGTCAATGAGCTTAAACAGGCAAGCTTCAATATCACTGCTTTGG 1620

QY 1621 ATGAAGGTTCTTTGTTGGCTCTGTGAGGGAAGACGTAAAGAGCCATGAGAGTCTG 1680  
DB 1621 ATGAAGGTTCTTTGTTGGCTCTGTGAGGGAAGACGTAAAGAGCCATGAGAGTCTG 1680

QY 1681 CTGGGCTGTCCCGTTCCTGGGGGTGAGAGAGGCTTGCAGTGGGTCTCTGTGTTG 1740  
DB 1681 CTGGGCTGTCCCGTTCCTGGGGGTGAGAGAGGCTTGCAGTGGGTCTCTGTGTTG 1740

QY 1741 GGTCAAGAGCTTAATGCAACCAACCCAGAGACACCTTGAAGGCTTCACTGTCTTTTC 1800  
DB 1741 GGTCAAGAGCTTAATGCAACCAACCCAGAGACACCTTGAAGGCTTCACTGTCTTTTC 1800

QY 1801 GAGACTCAAGACAAGAGTGTGCTGCTGCTGCTTAAACAGCTTCCAAAGAGTGTGCTT 1860  
DB 1801 GAGACTCAAGACAAGAGTGTGCTGCTGCTGCTTAAACAGCTTCCAAAGAGTGTGCTT 1860

QY 1861 CCGATTAACCAAGACTGAGCTTGAATGACATCTTCTGCTTCCAGCACTGTCCGTAT 1920  
DB 1861 CCGATTAACCAAGACTGAGCTTGAATGACATCTTCTGCTTCCAGCACTGTCCGTAT 1920

QY 1921 TTTGGGAAAAATTCGGGTGATGTCAAGAGATCTTCCCAAGATGAGTCCGCTGAGGCA 1980  
DB 1921 TTTGGGAAAAATTCGGGTGATGTCAAGAGATCTTCCCAAGATGAGTCCGCTGAGGCA 1980

QY 1981 TGTCTGTGTGCTCTTATGATGCGGGAATAGACCTCTCATTTGAGAGCAGTGGGAAGAT 2040  
DB 1981 TGTCTGTGTGCTCTTATGATGCGGGAATAGACCTCTCATTTGAGAGCAGTGGGAAGAT 2040

QY 2041 TTTGCTCAATGCTTTGCAACCAACCACTGTGGAGGCTGAGCTGGGAGAGCATC 2100  
DB 2041 TTTGCTCAATGCTTTGCAACCAACCACTGTGGAGGCTGAGCTGGGAGAGCATC 2100

QY 2101 CTGACAGAGGGGCTATGAGAGCTGTGTGTCGAAGCTGAGGCACTCCACTGCAAGATA 2160

Db 2101 CTGACAGACGGCCATGAAAGCCTGTGTGCCAAGCTGAGGACCTCCACTGCAAGATA 2160  
Qy 2161 CAGACCTGATGTTAGAAATGACAGATTACCCGTGTGTGAGACCTCTGAGAAATC 2220  
Db 2161 CAGACCTGATGTTAGAAATGACAGATTACCCGTGTGTGAGACCTCTGAGAAATC 2220  
Qy 2221 GTTCATGGCCAAACCGTAACTTAAGATCCCTCAACTTGGAGGACCCACCTGAAGAAAG 2280  
Db 2221 GTTCATGGCCAAACCGTAACTTAAGATCCCTCAACTTGGAGGACCCACCTGAAGAAAG 2280  
Qy 2281 GATGTAAAGATGGCGTGTGAAGCTTAAACACCCAAATGTTTGTGTGAAGCTTTGAG 2340  
Db 2281 GATGTAAAGATGGCGTGTGAAGCTTAAACACCCAAATGTTTGTGTGAAGCTTTGAG 2340  
Qy 2341 CTGGAATGCTGTGGAATTTGACCCATGCTGTATCTTGAAGATCTCCAAATCTTTAGAC 2400  
Db 2341 CTGGAATGCTGTGGAATTTGACCCATGCTGTATCTTGAAGATCTCCAAATCTTTAGAC 2400  
Qy 2401 TCCCCAGCCTGAATCTCTGAGCCTGGAGAAACAAAGTGAAGACAGACGAGAGTATG 2460  
Db 2401 TCCCCAGCCTGAATCTCTGAGCCTGGAGAAACAAAGTGAAGACAGACGAGAGTATG 2460  
Qy 2461 CCTCTAGATGATGCTTGAAGAGCTTCCAGTGGCCCTGCAAGAGCTGATCTGAGAGAC 2520  
Db 2461 CCTCTAGATGATGCTTGAAGAGCTTCCAGTGGCCCTGCAAGAGCTGATCTGAGAGAC 2520  
Qy 2521 TGTGGATCAAGCCACGGGTTGGCAGAGTCTGGCTCAGCCCTGTCAGCAACCGAGC 2580  
Db 2521 TGTGGATCAAGCCACGGGTTGGCAGAGTCTGGCTCAGCCCTGTCAGCAACCGAGC 2580  
Qy 2581 TTGACACACCTGACCTATCCAAACACAGCTGGGGAGGAAAGGTAAATCTAATCTGT 2640  
Db 2581 TTGACACACCTGACCTATCCAAACACAGCTGGGGAGGAAAGGTAAATCTAATCTGT 2640  
Qy 2641 CGATCCATGAGGCTTCCCACTGATGCTGACAGAGCTGATGCTGATGCTGATGCTG 2700  
Db 2641 CGATCCATGAGGCTTCCCACTGATGCTGACAGAGCTGATGCTGATGCTGATGCTG 2700  
Qy 2701 GACACGGCTGGCTGTGTGTTTCTTGAAGCTTGGCTTATGAGGATCTGATGCTGAC 2760  
Db 2701 GACACGGCTGGCTGTGTGTTTCTTGAAGCTTGGCTTATGAGGATCTGATGCTGAC 2760  
Qy 2761 CTGAGCCTTGAAGACCTGCTGAGAAACAAAGGTGAGCTTGTGCTGAGAGCTGATG 2820  
Db 2761 CTGAGCCTTGAAGACCTGCTGAGAAACAAAGGTGAGCTTGTGCTGAGAGCTGATG 2820  
Qy 2821 AGAGAACCATCTTGTATCTCCAGAGCTGAGAGTGGTAAAGTGTATCTCAACCGCGG 2880  
Db 2821 AGAGAACCATCTTGTATCTCCAGAGCTGAGAGTGGTAAAGTGTATCTCAACCGCGG 2880  
Qy 2881 TGTGTGAGAGTCTGTCTGTGTGATCTGAGAGAGACACCTGAGAGCTGATCTC 2940  
Db 2881 TGTGTGAGAGTCTGTCTGTGTGATCTGAGAGAGACACCTGAGAGCTGATCTC 2940  
Qy 2941 AGGACAAATGCGCTGGGTGAAGGTGGGTGCTGGCTGTGAGAGAGGATGAAGAAAG 3000  
Db 2941 AGGACAAATGCGCTGGGTGAAGGTGGGTGCTGGCTGTGAGAGAGGATGAAGAAAG 3000  
Qy 3001 AACAGTGTCTGACGAGCTCGGGGTGAAGCATGAGCTGATCTGATGCTGTGAG 3060  
Db 3001 AACAGTGTCTGACGAGCTCGGGGTGAAGCATGAGCTGATCTGATGCTGTGAG 3060  
Qy 3061 GCACTCTCTTGGCCCTTCTGCAACCGGACCTGACAGCTAAACCTGTGACGAAT 3120  
Db 3061 GCACTCTCTTGGCCCTTCTGCAACCGGACCTGACAGCTAAACCTGTGACGAAT 3120  
Qy 3121 AACCTCAGTCCCAAGAAATGAAGAGCTGTGTGGCTTGGCTGTGCTCCACGCTAAC 3180  
Db 3121 AACCTCAGTCCCAAGAAATGAAGAGCTGTGTGGCTTGGCTGTGCTCCACGCTAAC 3180  
Qy 3181 TTACAGATTAATGGCTGTGAAGATGACAGTACCTGCTGCAATAAGAAAGCTGTGAG 3240  
Db 3181 TTACAGATTAATGGCTGTGAAGATGACAGTACCTGCTGCAATAAGAAAGCTGTGAG 3240

Qy 3241 GAAATGACGCTTCTGAGCCCGAGTGGTAATTTGACGGATGTTGACATTTCTTTATGA 3300  
Db 3241 GAAATGACGCTTCTGAGCCCGAGTGGTAATTTGACGGATGTTGACATTTCTTTATGA 3300  
Qy 3301 GATGACCGACCAAAATAGACTTCTTCCGGCTCCCTGAAGCGGACATGACATG 3360  
Db 3301 GATGACCGACCAAAATAGACTTCTTCCGGCTCCCTGAAGCGGACATGACATG 3360  
Qy 3361 GCTTGTGTGGGAGTGAACCCAGAGCAAGAAAGCTGTGTGCTTGTGCTGAGAC 3420  
Db 3361 GCTTGTGTGGGAGTGAACCCAGAGCAAGAAAGCTGTGTGCTTGTGCTGAGAC 3420  
Qy 3421 TTCAAGACAGTACACGATTTGGCCAAAGTCTCTGCTGAGCCACCGCAATGATGCC 3480  
Db 3421 TTCAAGACAGTACACGATTTGGCCAAAGTCTCTGCTGAGCCACCGCAATGATGCC 3480  
Qy 3481 CAGAGATTTGACAAAGTGAAGAGAGCTCCCGCAACCATGAGGACAGGACACAA 3540  
Db 3481 CAGAGATTTGACAAAGTGAAGAGAGCTCCCGCAACCATGAGGACAGGACACAA 3540  
Qy 3541 CAAAGTAAATGTTGATGATGATATTTCCGAGCCTGTGAAATGCTGAGCTGAA 3600  
Db 3541 CAAAGTAAATGTTGATGATGATATTTCCGAGCCTGTGAAATGCTGAGCTGAA 3600  
Qy 3601 GGGCTTGAATCCAAAGTGTGATGATGACCAAGAGGATATGCTGTCTGATGAGAG 3660  
Db 3601 GGGCTTGAATCCAAAGTGTGATGATGACCAAGAGGATATGCTGTCTGATGAGAG 3660  
Qy 3661 GAGCTGAGCTGAGGGGCTTGTGTCACAGTGTGATGACCAAGCGGATGTCTGT 3720  
Db 3661 GAGCTGAGCTGAGGGGCTTGTGTCACAGTGTGATGACCAAGCGGATGTCTGT 3720  
Qy 3721 CACTGGAGCGGCTGGCTCTAGGGCTGTGTCTTAAACAGTGTGATGACCAAGCGT 3780  
Db 3721 CACTGGAGCGGCTGGCTCTAGGGCTGTGTCTTAAACAGTGTGATGACCAAGCGT 3780  
Qy 3781 GTGTCTGTGATCACTGGAGCGGCTGGCTCGAGGGCTTGTCTCAACAGTGTGATAC 3840  
Db 3781 GTGTCTGTGATCACTGGAGCGGCTGGCTCGAGGGCTTGTCTCAACAGTGTGATAC 3840  
Qy 3841 CACAGCGGTGAGCTGTGATCACTGGAGCGGCTGGCTCGAGGGCTTGTCTCAACAGT 3900  
Db 3841 CACAGCGGTGAGCTGTGATCACTGGAGCGGCTGGCTCGAGGGCTTGTCTCAACAGT 3900  
Qy 3901 GCTGATGACCAAGCGGTGTGTCTGTGATCACTGGAGCGGCTGGCTCGAGGGCTTGTG 3960  
Db 3901 GCTGATGACCAAGCGGTGTGTCTGTGATCACTGGAGCGGCTGGCTCGAGGGCTTGTG 3960  
Qy 3961 TCCAAACATGCTGATGACCAAGCGGTGTGTCTGTGATCACTGGAGCGGCTGGCTCGAG 4020  
Db 3961 TCCAAACATGCTGATGACCAAGCGGTGTGTCTGTGATCACTGGAGCGGCTGGCTCGAG 4020  
Qy 4021 GGGCTGTGTCTTAA 4035  
Db 4021 GGGCTGTGTCTTAA 4035

RESULT 2  
US-10-066-521-5  
; Sequence 5, Application US/1006521  
; Publication No. US20030027757A1  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; APPLICANT: Wang, Weiye  
; APPLICANT: Blatcher, Maria  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR  
; FILE REFERENCE: 07334-334001  
; CURRENT APPLICATION NUMBER: US/10/066, 521  
; CURRENT FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: 60/318, 645  
; PRIOR FILING DATE: 2001-09-10

PRIOR APPLICATION NUMBER: 60/265,231  
PRIOR FILING DATE: 2001-01-31  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 4035  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(4032)  
US-10-066-521-5.

Query Match 100.0%; Score 4035; DB 14; Length 4035;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAGAGCAAAATGCTCACTTTTCCAGCTACGGGCTGCATGSGTGTCTATGAG 60  
DB 1 ATGGAAGAGCAAAATGCTCACTTTTCCAGCTACGGGCTGCATGSGTGTCTATGAG 60  
QY 61 CTAGCAAGAGAAATTTCAACATTTCAAGAAATTAATAAGAAATTTCTCAGAAATCG 120  
DB 61 CTAGCAAGAGAAATTTCAACATTTCAAGAAATTAATAAGAAATTTCTCAGAAATCG 120  
QY 121 ACCACATGCTCTATTCACACAGTTTGAATCGAATGCCAAGTGAATGTCTGGCACTC 180  
DB 121 ACCACATGCTCTATTCACACAGTTTGAATTCAGAAATGCCAAGTGAATGTCTGGCACTC 180  
QY 181 CTCTTGATGATATTAATGAGATCGCTGCGCTGGGCTACGTCATTAGCATCTTTGAA 240  
DB 181 CTCTTGATGATATTAATGAGATCGCTGCGCTGGGCTACGTCATTAGCATCTTTGAA 240  
QY 241 AACATGAACCTGCGCAACCTCTCGGAGAAAGGCACGGGATGACATGAATAAAATTTCAAA 300  
DB 241 AACATGAACCTGCGCAACCTCTCGGAGAAAGGCACGGGATGACATGAATAAAATTTCAAA 300  
QY 301 GCTATGAAACAAGAGGTGCGCACAGCAGACAGACAGAAAGCAAGAAATTTCAACAGCT 360  
DB 301 GCTATGAAACAAGAGGTGCGCACAGCAGACAGACAGAAAGCAAGAAATTTCAACAGCT 360  
QY 361 ATGGAACAAGAGGTGCGCACAGCAGACAGACAGAAAGCAAGAAATGATGAGGTGACA 420  
DB 361 ATGGAACAAGAGGTGCGCACAGCAGACAGACAGAAAGCAAGAAATGATGAGGTGACA 420  
QY 421 TGGGACTCAAGAGTCAAGTATGACAAATTCGTGAGAGAGAGATGTAAGTCTGTAGT 480  
DB 421 TGGGACTCAAGAGTCAAGTATGACAAATTCGTGAGAGAGAGATGTAAGTCTGTAGT 480  
QY 481 TTTGAAAACACTGCTGCTGACTGAGCTGGCCGGAAATGCAAACTTGGCTTTTGATTCA 540  
DB 481 TTTGAAAACACTGCTGCTGACTGAGCTGGCCGGAAATGCAAACTTGGCTTTTGATTCA 540  
QY 541 GACCGGTGGGCTTCCGGCTCGACAGGTGTTCTGACGGAAGTCAAGAAATTTGGGAAA 600  
DB 541 GACCGGTGGGCTTCCGGCTCGACAGGTGTTCTGACGGAAGTCAAGAAATTTGGGAAA 600  
QY 601 TCGGCTCTAGCCAGAAGATGCTGCTGCTGGGCGCAAGTGAATCTTACCAAGGAATG 660  
DB 601 TCGGCTCTAGCCAGAAGATGCTGCTGCTGGGCGCAAGTGAATCTTACCAAGGAATG 660  
QY 661 TTTCTCTAGCTCTTTCTCTCCCGTTAGAGATGACGCGGAAGAGAGAGAGAGCTGTC 720  
DB 661 TTTCTCTAGCTCTTTCTCTCCCGTTAGAGATGACGCGGAAGAGAGAGAGAGCTGTC 720  
QY 721 ACAAGTTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGCTGACGGAATCATGTCC 780  
DB 721 ACAAGTTCATCTCCAGGAGTGGCCAGACTCCAGGCTCCGCTGACGGAATCATGTCC 780  
QY 781 CGACCAAGAAAGGCTGTGTTCAATTAAGACGTTTCAATGACCTGGGCTCTGTCTCAAC 840  
DB 781 CGACCAAGAAAGGCTGTGTTCAATTAAGACGTTTCAATGACCTGGGCTCTGTCTCAAC 840

QY 841 AATGACCAAAAGCTCTGCAAAAGACTGGGCTGAGAAAGAGCTCCGTTCACTCATATACG 900  
DB 841 AATGACCAAAAGCTCTGCAAAAGACTGGGCTGAGAAAGAGCTCCGTTCACTCATATACG 900  
QY 901 AGTCTGCTGAGAAAGGTCTGCTCTCTGAGTCTTCTGATTCGTACACCTGACAGAGAGTG 960  
DB 901 AGTCTGCTGAGAAAGGTCTGCTCTCTGAGTCTTCTGATTCGTACACCTGACAGAGAGTG 960  
QY 961 GGCACAGAGAAAGCTCAAGTCAAGAGGTGCTGCTCCCGTTACCTGTTAGTTAGAGAAATC 1020  
DB 961 GGCACAGAGAAAGCTCAAGTCAAGAGGTGCTGCTCCCGTTACCTGTTAGTTAGAGAAATC 1020  
QY 1021 TCCGGGAAACAAGAAATCACTTGCTCTTGAAGCGCGGATGAGTGAATCAGAGACA 1080  
DB 1021 TCCGGGAAACAAGAAATCACTTGCTCTTGAAGCGCGGATGAGTGAATCAGAGACA 1080  
QY 1081 CAAGGTTGGCTGATCATGAAACACCGTGAAGTCTCGACCAAGTGCAGAGTCCCGCC 1140  
DB 1081 CAAGGTTGGCTGATCATGAAACACCGTGAAGTCTCGACCAAGTGCAGAGTCCCGCC 1140  
QY 1141 GTGGGCTCTCATCTGCGTGGCCCTGACGCTGCAGAGACGTTGGGAGAGAGGCTGCGC 1200  
DB 1141 GTGGGCTCTCATCTGCGTGGCCCTGACGCTGCAGAGACGTTGGGAGAGAGGCTGCGC 1200  
QY 1201 CCTTCAACCAACGCTCAAGGCTGACAGCGCTTTTGTGTTCAATCAAGCTCACCCCT 1260  
DB 1201 CCTTCAACCAACGCTCAAGGCTGACAGCGCTTTTGTGTTCAATCAAGCTCACCCCT 1260  
QY 1261 CGAGGCTGTGCTCGGCGCTGCTCTCAATCTGAGAGAAAGTTGCTCTGAAGCGTTCTGC 1320  
DB 1261 CGAGGCTGTGCTCGGCGCTGCTCTCAATCTGAGAGAAAGTTGCTCTGAAGCGTTCTGC 1320  
QY 1321 CGTATGCTGTGAGGAGGTGGAATGAGAAAGTCAAGTTTGAATGATGAGTGAAGTCAATG 1380  
DB 1321 CGTATGCTGTGAGGAGGTGGAATGAGAAAGTCAAGTTTGAATGATGAGTGAAGTCAATG 1380  
QY 1381 GTTCAAGAGACTCGGAGAGTCTGAGCTCGCTCTGTTTCAATGAATCTCTTCCCA 1440  
DB 1381 GTTCAAGAGACTCGGAGAGTCTGAGCTCGCTCTGTTTCAATGAATCTCTTCCCA 1440  
QY 1441 GACAGCCACTGTGAGAGTACTACACTTCTTCACTCAAGTCTCCAGACTTCTGCGC 1500  
DB 1441 GACAGCCACTGTGAGAGTACTACACTTCTTCACTCAAGTCTCCAGACTTCTGCGC 1500  
QY 1501 GCTTTGACTACGTTTGAAGGAGCTGGAATCGAGCAGCTCTGCTCTGTAGCTT 1560  
DB 1501 GCTTTGACTACGTTTGAAGGAGCTGGAATCGAGCAGCTCTGCTCTGTAGCTT 1560  
QY 1561 GAGAAACAAGAGGTCATGAGACTTAAACAGGCAAGCTTCAATCACTGCTTTGG 1620  
DB 1561 GAGAAACAAGAGGTCATGAGACTTAAACAGGCAAGCTTCAATCACTGCTTTGG 1620  
QY 1621 ATGAAGGTTTCTTGTGCTCTGTAAGAGAGAGTAAAGAGGCACTGAGAGTCTTG 1680  
DB 1621 ATGAAGGTTTCTTGTGCTCTGTAAGAGAGAGTAAAGAGGCACTGAGAGTCTTG 1680  
QY 1681 CTGGGCTGTCCCGTTCCTGGGGGTGAAGAGAGCTTTCGACCTGGGTCTCTGTG 1740  
DB 1681 CTGGGCTGTCCCGTTCCTGGGGGTGAAGAGAGCTTTCGACCTGGGTCTCTGTG 1740  
QY 1741 GGTGAGAGCTTAATGCAACCCAGAGAGACCTTGAAGCGCTTCACTGTCTTTTC 1800  
DB 1741 GGTGAGAGCTTAATGCAACCCAGAGAGACCTTGAAGCGCTTCACTGTCTTTTC 1800  
QY 1801 GAGACTCAAGCAAAAGATTGTTGCTTGGCATTTAAACAGCTTCAAGAAAGTGTGCTT 1860  
DB 1801 GAGACTCAAGCAAAAGATTGTTGCTTGGCATTTAAACAGCTTCAAGAAAGTGTGCTT 1860  
QY 1861 CCGATTAAACAGAACTGGAATTGATGACATCTTCTGCTCCAGCACTGTCCGAT 1920  
DB 1861 CCGATTAAACAGAACTGGAATTGATGACATCTTCTGCTCCAGCACTGTCCGAT 1920  
QY 1921 TTGCGAAAAATTCGGGTGATGTCAAGAGGATCTTCCAAAGATGATGCTCGTAGGCA 1980



Db 1921 TTGGGAAAAATTCGGGTGATGTCAAGGATCTTCCAAAGATGATGAGTCCCTGAGGCA 1980  
 Qy 1981 TGTCTGTGTGTCCTCTATGATGTGGGGATTAAGACCTCATTTGAGAGCATGGGAAAGAT 2040  
 Db 1981 TGTCTGTGTGTCCTCTATGATGTGGGGATTAAGACCTCATTTGAGAGCATGGGAAAGAT 2040  
 Qy 2041 TTCTGTCCATGCTTGGGACCCAGCCAGACCTGGGGAGCTGGACCTGGGGAGCAGCATC 2100  
 Db 2041 TTCTGTCCATGCTTGGGACCCAGCCAGACCTGGGGAGCTGGACCTGGGGAGCAGCATC 2100  
 Qy 2101 CTGACAGAGCGGGCCATGAAGACCTGTGTGCAAGCTGAGGCACTCCACTGCAAGATA 2160  
 Db 2101 CTGACAGAGCGGGCCATGAAGACCTGTGTGCAAGCTGAGGCACTCCACTGCAAGATA 2160  
 Qy 2161 CAGACCCGTGATGTTTGAATGCAAGATTACCCCTGTGTGAGCAGCATCTTGAGAAATC 2220  
 Db 2161 CAGACCCGTGATGTTTGAATGCAAGATTACCCCTGTGTGAGCAGCATCTTGAGAAATC 2220  
 Qy 2221 GTCATGGCCAAACGTTAAGATCCCTCAACTTGGGAGGAGCCACTGAAGAGAG 2280  
 Db 2221 GTCATGGCCAAACGTTAAGATCCCTCAACTTGGGAGGAGCCACTGAAGAGAG 2280  
 Qy 2281 GATGTAAAGATGCGGTGTGAAGCCTTAAACACCCAAATGTTGTGAGTCTTTGAGG 2340  
 Db 2281 GATGTAAAGATGCGGTGTGAAGCCTTAAACACCCAAATGTTGTGAGTCTTTGAGG 2340  
 Qy 2341 CTGGAATGTGTGATTTGACCCATGCTGTAACTGTGAAGATCTCCAAATCTTACGACC 2400  
 Db 2341 CTGGAATGTGTGATTTGACCCATGCTGTAACTGTGAAGATCTCCAAATCTTACGACC 2400  
 Qy 2401 TCCCCAGGCTTAAATCTTGAGCCTGGGAGAGAAACAAGTGAACAAGAGAGATATG 2460  
 Db 2401 TCCCCAGGCTTAAATCTTGAGCCTGGGAGAGAAACAAGTGAACAAGAGAGATATG 2460  
 Qy 2461 CCTCTAGTATGCTTGAAGTCTCCAGTGGGCGCTGACAGAGCTGTAATCTGAGAGAC 2520  
 Db 2461 CCTCTAGTATGCTTGAAGTCTCCAGTGGGCGCTGACAGAGCTGTAATCTGAGAGAC 2520  
 Qy 2521 TGTGGCATCAAGCCAGCGGTTGCCAGAGTCTGGCCTCAGCCCTGTGACAAACCGAGAC 2580  
 Db 2521 TGTGGCATCAAGCCAGCGGTTGCCAGAGTCTGGCCTCAGCCCTGTGACAAACCGAGAC 2580  
 Qy 2581 TTGAACACACTGTGCTATCCAAACAAGCTGGGGAAAGAGGTGAATCTACTGTGT 2640  
 Db 2581 TTGAACACACTGTGCTATCCAAACAAGCTGGGGAAAGAGGTGAATCTACTGTGT 2640  
 Qy 2641 CGATCCATGAGGCTTCCCACTGTAGTCTGCAAGAGGCTGATCTGAATCAGTCCACTG 2700  
 Db 2641 CGATCCATGAGGCTTCCCACTGTAGTCTGCAAGAGGCTGATCTGAATCAGTCCACTG 2700  
 Qy 2701 GACACGGCTGCTGTGTGTTTCTTGCACCTTGGCTTAATGGATCACTAGGCTGACGAC 2760  
 Db 2701 GACACGGCTGCTGTGTGTTTCTTGCACCTTGGCTTAATGGATCACTAGGCTGACGAC 2760  
 Qy 2761 CTGAGCCTTAAGATGAACCTGTGGAAGAACAATGGCTGGAAGCTTGTGGAGGT CATG 2820  
 Db 2761 CTGAGCCTTAAGATGAACCTGTGGAAGAACAATGGCTGGAAGCTTGTGGAGGT CATG 2820  
 Qy 2821 AGAGAACAATCTTGCATCTCAAGACCTGGAGTTGTAAAGTGCATCTCACCGCCGG 2880  
 Db 2821 AGAGAACAATCTTGCATCTCAAGACCTGGAGTTGTAAAGTGCATCTCACCGCCGG 2880  
 Qy 2881 TGTCTGTGAGAGTCTGTCTGTGTGATCTTGAAGAGCAGACCTGAAGAGCTGATCTC 2940  
 Db 2881 TGTCTGTGAGAGTCTGTCTGTGTGATCTTGAAGAGCAGACCTGAAGAGCTGATCTC 2940  
 Qy 2941 ACGGACAATGCTTGGGTGAACGCTGGGGTGTCTGGCTGTGCGAGGAGCTGAAGCAAAAG 3000  
 Db 2941 ACGGACAATGCTTGGGTGAACGCTGGGGTGTCTGGCTGTGCGAGGAGCTGAAGCAAAAG 3000  
 Qy 3001 AACAGTGTCTGACGAGACTCGGGTTGAAGGCAATGTGACATCTTGATGTGCTGAG 3060  
 Db 3001 AACAGTGTCTGACGAGACTCGGGTTGAAGGCAATGTGACATCTTGATGTGCTGAG 3060

Db 3001 AACAGTGTCTGACGAGACTCGGGTTGAAGGCAATGTGACATCTTGATGTGCTGAG 3060  
 Qy 3061 GCACTCTCTTGGGCTTTCTGCAACCGGATCTGACCACTTAAACCTGTGAGAT 3120  
 Db 3061 GCACTCTCTTGGGCTTTCTGCAACCGGATCTGACCACTTAAACCTGTGTGAGAT 3120  
 Qy 3121 AACTTCAGTCCAAAGAAATGATGAAGCTGTGTGCGGCTTGGCTGTGCCAGCTAAC 3180  
 Db 3121 AACTTCAGTCCAAAGAAATGATGAAGCTGTGTGCGGCTTGGCTGTGCCAGCTAAC 3180  
 Qy 3181 TTACAGATTAATTTGGCTGTGAAATGAGCATCCCTGTGCAATTAAGAGCTGTGAG 3240  
 Db 3181 TTACAGATTAATTTGGCTGTGAAATGAGCATCCCTGTGCAATTAAGAGCTGTGAG 3240  
 Qy 3241 GAAGTGCAGCTTCTAACCCCGAGTGTGTAATTTGACGTAATTTGCTTTTATGA 3300  
 Db 3241 GAAGTGCAGCTTCTAACCCCGAGTGTGTAATTTGACGTAATTTGCTTTTATGA 3300  
 Qy 3301 GATGACCGACAATAATGAGCTTAATCTTCCGGCTCCCTGAAGCCGGGCAATGGCATGT 3360  
 Db 3301 GATGACCGACAATAATGAGCTTAATCTTCCGGCTCCCTGAAGCCGGGCAATGGCATGT 3360  
 Qy 3361 GCTTGTCTGTGGGATGAACCCAGAGAGAGAAAGCCTGTGTGCTTCTGGCTGAGAC 3420  
 Db 3361 GCTTGTCTGTGGGATGAACCCAGAGAGAGAAAGCCTGTGTGCTTCTGGCTGAGAC 3420  
 Qy 3421 TTCAAAGCATGACAGATTTGGCAAGTCTTCTGTGCTGAGCAAGCAATGTGATCC 3480  
 Db 3421 TTCAAAGCATGACAGATTTGGCAAGTCTTCTGTGCTGAGCAAGCAATGTGATCC 3480  
 Qy 3481 CAGAGAGTTGAACAATGTGAGCAAGAGCTCCCGGCAACCAAGGAGAGAGCAAGCA 3540  
 Db 3481 CAGAGAGTTGAACAATGTGAGCAAGAGCTCCCGGCAACCAAGGAGAGAGCAAGCA 3540  
 Qy 3541 CAAGATTAATAATTTGAGTGTGATTAATCCGAGGCTGTCTGAATCTGAGCTGAA 3600  
 Db 3541 CAAGATTAATAATTTGAGTGTGATTAATCCGAGGCTGTGTGAATCTGAGCTGAA 3600  
 Qy 3601 GGGCTTGATCAACAGTGTGATCAATGACCAAGAGGATAGCCTGTCTACTAGAGAG 3660  
 Db 3601 GGGCTTGATCAACAGTGTGATCAATGACCAAGAGGATAGCCTGTCTACTAGAGAG 3660  
 Qy 3661 GAGCTGACCTGAGGGGCTTGTGTCACAAGTGTGATGAACAAGCGGTGTGCTGTGT 3720  
 Db 3661 GAGCTGACCTGAGGGGCTTGTGTCACAAGTGTGATGAACAAGCGGTGTGCTGTGT 3720  
 Qy 3721 CACTGGAGCGGCTGTGAGGCTGTGTCTTAACAGTGTGATGAACAAGCGGT 3780  
 Db 3721 CACTGGAGCGGCTGTGAGGCTGTGTCTTAACAGTGTGATGAACAAGCGGT 3780  
 Qy 3781 GTGTCTGTGATCAATGGAGCGGCTGTGAGGCTGTGTCTTAACAGTGTGATGAAC 3840  
 Db 3781 GTGTCTGTGATCAATGGAGCGGCTGTGAGGCTGTGTCTTAACAGTGTGATGAAC 3840  
 Qy 3841 CACAGCGGTGTGCTGTGATCAATGGAGCGGCTGTGAGGCTGTGTCTTAACAGT 3900  
 Db 3841 CACAGCGGTGTGCTGTGATCAATGGAGCGGCTGTGAGGCTGTGTCTTAACAGT 3900  
 Qy 3901 GCTGATGACCAAGCGGTGTGCTGTGATCAATGGAGCGGCTGTGAGGCTGTGTG 3960  
 Db 3901 GCTGATGACCAAGCGGTGTGCTGTGATCAATGGAGCGGCTGTGAGGCTGTGTG 3960  
 Qy 3961 TCCAAACAGTGTGATGAACAAGCGGTGTGCTGTGATCAATGGAGCGGCTGTGAG 4020  
 Db 3961 TCCAAACAGTGTGATGAACAAGCGGTGTGCTGTGATCAATGGAGCGGCTGTGAG 4020  
 Qy 4021 GGGCTGTGTCTTAA 4035  
 Db 4021 GGGCTGTGTCTTAA 4035

Sequence 3, Application US/10416642  
Publication No. US20040043452A1  
GENERAL INFORMATION:  
APPLICANT: INCYTE GENOMICS, INC.  
APPLICANT: RAMKUMAR, Jayalaxmi  
APPLICANT: ARVIZU, Chandra  
TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS  
FILE REFERENCE: PF-0842 PCT  
CURRENT APPLICATION NUMBER: US/10/416,642  
PRIOR FILING DATE: 2003-05-13  
PRIOR APPLICATION NUMBER: 60/249,407  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PERL Program  
SEQ ID NO 3  
LENGTH: 3489  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20040043452A1 7474830CB1  
US-10-416-642-3

Query Match 72.3%; Score 2917; DB 18; Length 3489;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 290 AAATTTCAGCTATGGAACAAGAGGTGCGACAGCAGACAGACAGACAGAA 349  
DB 452 AAATTTCAGCTATGGAACAAGAGGTGCGACAGCAGACAGACAGACAGAA 511  
QY 350 TTTTCAAGCTATGGAACAAGAGGTGCGACAGCAGACAGACAGACAGAA 409  
DB 512 TTTTCAAGCTATGGAACAAGAGGTGCGACAGCAGACAGACAGACAGAA 571  
QY 410 GAGGTGACATGAGGACTCAAGAGTCACTGATGACCAAAATTCGTTGAGAGAGATG 469  
DB 572 GAGGTGACATGAGGACTCAAGAGTCACTGATGACCAAAATTCGTTGAGAGAGATG 631  
QY 470 TACGTCGATGTTTGAACAACCTGCTGCTGACAGCGCGGAAATGAAACGTTGCTGCTG 529  
DB 632 TACGTCGATGTTTGAACAACCTGCTGCTGACAGCGCGGAAATGAAACGTTGCTGCTG 691  
QY 530 CTTTGTATTCAGACCGGTGAGGCTTCCGAGCTCGACAGGTGATCTGACAGAAAGTCA 589  
DB 692 CTTTGTATTCAGACCGGTGAGGCTTCCGAGCTCGACAGGTGATCTGACAGAAAGTCA 751  
QY 590 GAATTGGGAATTCGCTTACGACAGAGAGATGCTGCTGCTGAGGCGCAAGGTGACTCT 649  
DB 752 GAATTGGGAATTCGCTTACGACAGAGAGATGCTGCTGCTGAGGCGCAAGGTGACTCT 811  
QY 650 ACCGAGGAATGTTCCCTACGCTCTTCCCTCCCGTTAGAGAGATGAGGGAAGAG 709  
DB 812 ACCGAGGAATGTTCCCTACGCTCTTCCCTCCCGTTAGAGAGATGAGGGAAGAG 871  
QY 710 AGAGAGGTGACAGAGTTCATCTCCAGAGGTGCGACAGCTCCAGGCTCCGATGACG 769  
DB 872 AGAGAGGTGACAGAGTTCATCTCCAGAGGTGCGACAGCTCCAGGCTCCGATGACG 931  
QY 770 AGATCATGTCCGACCAAGAGGCTGTTGTTTCATTTGACGATTCATGACCTGAGCT 829  
DB 932 AGATCATGTCCGACCAAGAGGCTGTTGTTTCATTTGACGATTCATGACCTGAGCT 991  
QY 830 CTGTCCTCAAGACAGACCAAGCTCTGCAAGAGATGAGGCTGAGAGAGAGCTCCGTTCA 889  
DB 992 CTGTCCTCAAGACAGACCAAGCTCTGCAAGAGATGAGGCTGAGAGAGAGCTCCGTTCA 1051  
QY 890 CCTCATACGAGAGTGTGCTGAGAGAGGTCTGCTCCCTGAGAGCTCTTCTGATCTGACCG 949  
DB 1052 CCTCATACGAGAGTGTGCTGAGAGAGGTCTGCTCCCTGAGAGCTCTTCTGATCTGACCG 1111  
QY 950 TCAGAGAGCTGGGACAGAGAGCTCAAGTCAAGAGTCTGCTCTCCGTTACCTGTTAG 1009

DB 1112 TCAGAGAGCTGGGACAGAGAGCTCAAGTCAAGAGTCTGCTCTCCGTTACCTGTTAG 1171  
QY 1010 TTAGAGGAATCTCCGGGGAACAAGATTCACCTTGCTCTTGAAGCGGGATTGATGAGC 1069  
DB 1172 TTAGAGGAATCTCCGGGGAACAAGATTCACCTTGCTCTTGAAGCGGGATTGATGAGC 1231  
QY 1070 ATCAGAGACACAAAGGTTGCTGCTGATCAAGACAAAGTCAAGTCTGACCAAGTGC 1129  
DB 1232 ATCAGAGACACAAAGGTTGCTGCTGATCAAGACAAAGTCAAGTCTGACCAAGTGC 1291  
QY 1130 AGGTGCTCCGCTGAGGCTCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1189  
DB 1292 AGGTGCTCCGCTGAGGCTCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1351  
QY 1190 AGAGAGTGGCCCTTCAACCAAGCTCAAGGCTGACAGGCTGCTGCTGCTGCTGCTGCTG 1249  
DB 1352 AGAGAGTGGCCCTTCAACCAAGCTCAAGGCTGACAGGCTGCTGCTGCTGCTGCTGCTG 1411  
QY 1250 AGCTCACCCCTCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1309  
DB 1412 AGCTCACCCCTCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1471  
QY 1310 AGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1369  
DB 1472 AGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1531  
QY 1370 ACAGACTCATAGTTCAGAGACTCGGGAGTCTGAGCTCCGCTGCTGCTGCTGCTGCTGCTG 1429  
DB 1532 ACAGACTCATAGTTCAGAGACTCGGGAGTCTGAGCTCCGCTGCTGCTGCTGCTGCTGCTG 1591  
QY 1430 TCCCTTCTCCAGACAGCACTGTGAGAGTACTACACTTCTTCCACTGACTGCTCAGG 1489  
DB 1592 TCCCTTCTCCAGACAGCACTGTGAGAGTACTACACTTCTTCCACTGACTGCTCAGG 1651  
QY 1490 ACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1549  
DB 1652 ACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1711  
QY 1550 CTCTGTAAGTTCAGAGACAAAGAGTCAATGAGCTTTAAACAGGAGCTTCCATATCC 1609  
DB 1712 CTCTGTAAGTTCAGAGACAAAGAGTCAATGAGCTTTAAACAGGAGCTTCCATATCC 1771  
QY 1610 ACTGCTTTGATGATGAGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1669  
DB 1772 ACTGCTTTGATGATGAGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1831  
QY 1670 TGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1729  
DB 1832 TGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1891  
QY 1730 TCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1789  
DB 1892 TCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1951  
QY 1790 ACTGCTTTTGAAGCTCAAGACAAAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1849  
DB 1952 ACTGCTTTTGAAGCTCAAGACAAAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2011  
QY 1850 AAGTGTGCTTCGATTAACCAAGACTGATGATGATGATGATGATGATGATGATGATGATG 1909  
DB 2012 AAGTGTGCTTCGATTAACCAAGACTGATGATGATGATGATGATGATGATGATGATGATG 2071  
QY 1910 ACTGTCGTAATTTGCGGAAATTTGAGGTGATGTCAAAGGATTTTCCCAAGAGATGAGT 1969  
DB 2072 ACTGTCGTAATTTGCGGAAATTTGAGGTGATGTCAAAGGATTTTCCCAAGAGATGAGT 2131  
QY 1970 CCGCTGAGGCAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2029  
DB 2132 CCGCTGAGGCAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2191  
QY 2030 AGTGGGAAGATTTTCTGCTCAATGCTTGGCACCCACACCTGCGGAGTGGACCTGCG 2089  
DB 2192 AGTGGGAAGATTTTCTGCTCAATGCTTGGCACCCACACCTGCGGAGTGGACCTGCG 2251

2090 GCAGCAGCATCTCTGACAGAGCGGGCCATGAAGACCTGTGTGCGAAGCTGAGGCATCCCA 2149  
2232 GCAGCAGCATCTCTGACAGAGCGGGCCATGAAGACCTGTGTGCGAAGCTGAGGCATCCCA 2311  
2150 CCTGCAAGATACAGACCTGTATGTTTAAAGATGACAGATTACCCCTGTGTGCGACACC 2209  
2312 CCTGCAAGATACAGACCTGTATGTTTAAAGATGACAGATTACCCCTGTGTGCGACACC 2371  
2210 TCTGAGAAATCGTCATGCGCAACCGTAACCTTAAGATCCCTCAACTTTGGAGGACCCACC 2269  
2372 TCTGAGAAATCGTCATGCGCAACCGTAACCTTAAGATCCCTCAACTTTGGAGGACCCACC 2431  
2270 TGAAGAGAGAGATGTAAGAGATGGGTGGAAGCCCTTAAACACCCAAATGTTGTTGG 2329  
2432 TGAAGAGAGAGATGTAAGAGATGGGTGGAAGCCCTTAAACACCCAAATGTTGTTGG 2491  
2330 AGTCTTTGAGGCTGATGCTGTGTGATGACCCATGCTGTTACCTGAAGATCTCCCAA 2389  
2492 AGTCTTTGAGGCTGATGCTGTGTGATGACCCATGCTGTTACCTGAAGATCTCCCAA 2551  
2390 TCTTTACGACCTCTCCCGACCTGAAATCTTGAGCCTGGCAGAGAAACAAGTGACAGACC 2449  
2552 TCTTTACGACCTCTCCCGACCTGAAATCTTGAGCCTGGCAGAGAAACAAGTGACAGACC 2611  
2450 AGGAGTAATGCTCTCACTGATGCTGTAAGAGTCTCCAGTGGCCCTGGCAGAGATCTGA 2509  
2612 AGGAGTAATGCTCTCACTGATGCTGTAAGAGTCTCCAGTGGCCCTGGCAGAGATCTGA 2671  
2510 TACTGAGAGACTGTGGCATCAAGCCAGGCTGAGAGTGTGGCTCAGCCCTGCTCA 2569  
2672 TACTGAGAGACTGTGGCATCAAGCCAGGCTGAGAGTGTGGCTCAGCCCTGCTCA 2731  
2570 GCAACCGAGCTTGAACAACCTGTGCTTATCCAAACAAGCCTGGGAGAAACGAAGTGTAA 2629  
2732 GCAACCGAGCTTGAACAACCTGTGCTTATCCAAACAAGCCTGGGAGAAACGAAGTGTAA 2791  
2650 ATCTACTGTGTGATGCAAGAGGCTTCCCACTGTAAGTGTGAGAGGCTGATGCTGAATC 2689  
2792 ATCTACTGTGTGATGCAAGAGGCTTCCCACTGTAAGTGTGAGAGGCTGATGCTGAATC 2851  
2690 AGTGCCACTGTGACACAGGCTGTGTGTTTCTTGACCTTGAGGCTTATGAGTACTCAT 2749  
2852 AGTGCCACTGTGACACAGGCTGTGTGTTTCTTGACCTTGAGGCTTATGAGTACTCAT 2911  
2750 GAGTGACGACCTGAGCCTTACATGAACCTGTGAGAAACAAGTGTGAGTGTGTCTGT 2809  
2912 GAGTGACGACCTGAGCCTTACATGAACCTGTGAGAAACAAGTGTGAGTGTGTCTGT 2971  
2810 GCGAGGTATGAGAGAAACATTTGTTCATCTTCAGAGACTGTGAGTGTGTAAGTGTCTATC 2869  
2972 GCGAGGTATGAGAGAAACATTTGTTCATCTTCAGAGACTGTGAGTGTGTAAGTGTCTATC 3031  
2870 TCACGCGCGCGGTGTGTGAGAGTGTGTCTGTGTGATCTCGAGAGACGACACCTGAAGA 2929  
3032 TCACGCGCGCGGTGTGTGAGAGTGTGTCTGTGTGATCTCGAGAGACGACACCTGAAGA 3091  
2930 GCGTGTGATCTCAACGACATGCTGTGTGAGAGTGTGTGAGTGTGTGAGTGTGTGAGTGTG 2989  
3092 GCGTGTGATCTCAACGACATGCTGTGTGAGAGTGTGTGAGTGTGTGAGTGTGTGAGTGTG 3151  
2990 TGAAGCAAAAGAAAGAGTGTGTGAGAGAGAGTGTGTGAGAGAGTGTGTGAGTGTGTGAG 3049  
3152 TGAAGCAAAAGAAAGAGTGTGTGAGAGAGAGTGTGTGAGAGAGTGTGTGAGTGTGTGAG 3211  
3050 ATTGCTGTGAGGACCTCTCTTGTGCGCTTCTTCTGCAACCGGACCTGACAGTCTAAAC 3109  
3212 ATTGCTGTGAGGACCTCTCTTGTGCGCTTCTTCTGCAACCGGACCTGACAGTCTAAAC 3271  
3110 TGTGTGCAAGATTAATCTTCAAGTGTGCAAGAGATGATGATGATGATGATGATGATGATG 3169  
3272 TGTGTGCAAGATTAATCTTCAAGTGTGCAAGAGATGATGATGATGATGATGATGATGATG 3331

QY 3170 CCACGTCTAATCTACAGATTAATGAGCTGTGAAATGAGCAGTACCTGTGCAAAATGAAGA 3229  
DB 3332 CCACGTCTAATCTACAGATTAATGAGCTGTGAAATGAGCAGTACCTGTGCAAAATGAAGA 3391  
QY 3320 AGCTGTGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3289  
DB 3392 AGCTGTGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3451  
QY 3290 CTTTGTGATGAAGATGACCG 3308  
DB 3452 CTTTGTGATGAAGATGACCG 3470

## RESULT 4

US-10-216-645-1  
; Sequence 1, Application US/10216645  
; Publication No. US20030125282A1  
; GENERAL INFORMATION:  
; APPLICANT: WEISS, BERTRAM  
; APPLICANT: LESSL, MONIKA  
; APPLICANT: PETERS-KOTTIG, MICHAEL  
; APPLICANT: BECKMANN, GEORG  
; TITLE OF INVENTION: HUMAN MATER PROTEINS  
; FILE REFERENCE: SCH-1910  
; CURRENT APPLICATION NUMBER: US/10/216,645  
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 3926  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-216-645-1

## Query Match

Best Local Similarity 99.9%; Pred. No. 0; DB 15; Length 3926;  
Matches 3017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 290 AAATTTCAAGCTATGAGACAAAGAGTGCACAGCAGCAGACAGAGAAACAAGAA 349  
DB 452 AAATTTCAAGCTATGAGACAAAGAGTGCACAGCAGCAGACAGAGAAACAAGAA 511  
QY 350 TTTCAAGCTATGAGACAAAGAGTGCACAGCAGCAGACAGAGAAACAAGAA 409  
DB 512 TTTCAAGCTATGAGACAAAGAGTGCACAGCAGCAGACAGAGAAACAAGAA 571  
QY 410 GAGTGACACATGAGACATGAAGAGTCAAGTGAATGACCAATTTGCTGAGAGAGATG 469  
DB 572 GAGTGACACATGAGACATGAAGAGTCAAGTGAATGACCAATTTGCTGAGAGAGATG 631  
QY 470 TACGTGTGATTTGAAACACTGCTGTGACCTGAGCGGAGAAATCAAACTTTGCTGTG 529  
DB 632 TACGTGTGATTTGAAACACTGCTGTGACCTGAGCGGAGAAATCAAACTTTGCTGTG 691  
QY 530 CTTTGTGATGAGACCGTGTGGGCTTCCGCTGTGACGAGTGTCTGTGACGAGAAAGTCA 589  
DB 692 CTTTGTGATGAGACCGTGTGGGCTTCCGCTGTGACGAGTGTCTGTGACGAGAAAGTCA 751  
QY 590 GAAATTTGGAATTCGGCTCTTACCCAGAAAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 649  
DB 752 GAAATTTGGAATTCGGCTCTTACCCAGAAAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 811  
QY 650 ACCAGGAAATTTCTCTCAAGCTTTCTTCTCTCCCGTTAGAGATGACGAGGAGAAAG 709  
DB 812 ACCAGGAAATTTCTCTCAAGCTTTCTTCTCTCCCGTTAGAGATGACGAGGAGAAAG 871  
QY 710 AGAGCAAGTGTACAGAGTCTTCTCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 769  
DB 872 AGAGCAAGTGTACAGAGTCTTCTCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 931  
QY 770 AGATCATGTCCCGACAGAGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 829

Db 932 AGATCATGTCCCGACCAAGAAAGGCTGTTGTCATCATTTGACGGTTTCATGACCTGGGCT 991  
Qy 830 CTGTCTCTCAACATATACCAAAAGCTCTGCAAAAGATCTGGGCTGAGAAAGACCTCCGTTCA 889  
Db 992 CTGTCTCTCAACATATACCAAAAGCTCTGCAAAAGATCTGGGCTGAGAAAGACCTCCGTTCA 1051  
Qy 890 CCTCTCATAGCAGATGTGTGAGAAAGATCTGCTCCCTGAGATCTTCTCTGATTCGTCAACG 949  
Db 1052 CCTCTCATAGCAGATGTGTGAGAAAGATCTGCTCCCTGAGATCTTCTCTGATTCGTCAACG 1111  
Qy 950 TCAGAGAGCTGGGCA CAGAGAACTCAAGTCAGAGATCTGTCTTCCCGTTAACTCTTTAG 1009  
Db 1112 TCAGAGAGCTGGGCA CAGAGAACTCAAGTCAGAGATCTGTCTTCCCGTTAACTCTTTAG 1171  
Qy 1010 TTAAGAGAAATCTCCGGGGAA CAAAGATCCACTTCTCTTGAAGCGGGATTTGTTAGC 1069  
Db 1172 TTAAGAGAAATCTCCGGGGAA CAAAGATCCACTTCTCTTGAAGCGGGATTTGTTAGC 1231  
Qy 1070 ATCAGAGACA CAAAGGTTGGCGTGATCATGAACAA CCGTAGAGCTGCTCGACCACTGSC 1129  
Db 1232 ATCAGAGACA CAAAGGTTGGCGTGATCATGAACAA CCGTAGAGCTGCTCGACCACTGSC 1291  
Qy 1130 AGGTGCGCGCTGGGCTCTCTCATCTGCGCTGACCTGCACTGAGAGACGTGTGGGG 1189  
Db 1292 AGGTGCGCGCTGGGCTCTCTCATCTGCGCTGACCTGCACTGAGAGACGTGTGGGG 1351  
Qy 1190 AGAGGTGCGCGCTCTTCAACCAAAAGCTCA CAGGCTGCAAGCGCTTTTGTGTTATC 1249  
Db 1352 AGAGGTGCGCGCTCTTCAACCAAAAGCTCA CAGGCTGCAAGCGCTTTTGTGTTATC 1411  
Qy 1250 AGCTCAACCCCTGAGAGCGTGTGCTCGGCGCTGCTCAATCTGAGAGAAAGATGTCCTGA 1309  
Db 1412 AGCTCAACCCCTGAGAGCGTGTGCTCGGCGCTGCTCAATCTGAGAGAAAGATGTCCTGA 1471  
Qy 1310 AGCGCTTCTGCGTATGAGCTGTGAGAGAGTGTGAATAGAAAGTCAAGTGTGATGTTG 1369  
Db 1472 AGCGCTTCTGCGTATGAGCTGTGAGAGAGTGTGAATAGAAAGTCAAGTGTGATGTTG 1531  
Qy 1370 ACGACCTATGTTCAAGGACTCGGGGAGTCTGAGCTCCGCTCTGTTTCAANTGAACA 1429  
Db 1532 ACGACCTATGTTCAAGGACTCGGGGAGTCTGAGCTCCGCTCTGTTTCAANTGAACA 1591  
Qy 1430 TCTCTTCTCCAGACAGCACTGTGAGAGTACTACACTTCTTCCACTGAGTCTCCAG 1489  
Db 1592 TCTCTTCTCCAGACAGCACTGTGAGAGTACTACACTTCTTCCACTGAGTCTCCAG 1651  
Qy 1490 ACTTCTGTGCGCTTGTACTACGTGTTAGAGGCGTGAATTCGAGCCAGCTCTTGC 1549  
Db 1652 ACTTCTGTGCGCTTGTACTACGTGTTAGAGGCGTGAATTCGAGCCAGCTCTTGC 1711  
Qy 1550 CTCTGTAGTGAAGAA CAAAGAGTTCATGAGACTTAA CAGCAGAGCTTCCATATCC 1609  
Db 1712 CTCTGTAGTGAAGAA CAAAGAGTTCATGAGACTTAA CAGCAGAGCTTCCATATCC 1771  
Qy 1610 ACTGCTTTGATGAAGCGTTTCTGTTTGGCTGTGAGCGAAGACTTAAGAGAGCCAC 1669  
Db 1772 ACTGCTTTGATGAAGCGTTTCTGTTTGGCTGTGAGCGAAGACTTAAGAGAGCCAC 1831  
Qy 1670 TGAAGTCTGTGAGGCTGTCCCGTTCCCTGGGGGTGAAGCAAGAGCTTCTGACTGG 1729  
Db 1832 TGAAGTCTGTGAGGCTGTCCCGTTCCCTGGGGGTGAAGCAAGAGCTTCTGACTGG 1891  
Qy 1730 TCTCTCTGTGGGTCAAGAGCTTAATGCAACACCCAGAGACACCTTGAAGCTTCC 1789  
Db 1892 TCTCTCTGTGGGTCAAGAGCTTAATGCAACACCCAGAGACACCTTGAAGCTTCC 1951  
Qy 1790 ACTGCTTTTGAAGCTCAAGCAAAAGATTTGTTGCTTGGCACTTAAAGCTTCCAG 1849  
Db 1952 ACTGCTTTTGAAGCTCAAGCAAAAGATTTGTTGCTTGGCACTTAAAGCTTCCAG 2011  
Qy 1850 AAGTGTGCTTCCGATTAACAGAACTGGACTTGATGACTTCTTCTGCTCCAGC 1909

Db 2012 AAGTGTGCTTCCGATTAACAGAACTTGAAGACTTCTTCTGCTCCAGC 2071  
Qy 1910 ACTGTCCGTAATTTCCGGAATAATTCGGGTGATGTCAAAAGGATCTTCCAAAGATAGT 1969  
Db 2072 ACTGTCCGTAATTTCCGGAATAATTTGGGTGATGTCAAAAGGATCTTCCAAAGATAGT 2131  
Qy 1970 CCGCTGAGGCAATGTCTGTGTGCTTCTCTATGATGCGGGATTAAGACCTCATTTAGAGAC 2029  
Db 2132 CCGCTGAGGCAATGTCTGTGTGCTTCTCTATGATGCGGGATTAAGACCTCATTTAGAGAC 2191  
Qy 2030 AGTGGGAAGATTTCTGCTCCATGCTTTGGCACCCACCA CACTGCGGCACTGACTGG 2089  
Db 2192 AGTGGGAAGATTTCTGCTCCATGCTTTGGCACCCACCA CACTGCGGCACTGACTGG 2251  
Qy 2090 GCACAGATCTTCAAGAGGCGGCAATGAAGACCTGTGTGCCAAGCTGAGGATCCA 2149  
Db 2252 GCACAGATCTTCAAGAGGCGGCAATGAAGACCTGTGTGCCAAGCTGAGGATCCA 2311  
Qy 2150 CCTGCAAGATPACAGACCTGATGTTTGAATAATGACAGATTAACCTCTGTGTGACAGACC 2209  
Db 2312 CCTGCAAGATPACAGACCTGATGTTTGAATAATGACAGATTAACCTCTGTGTGACAGACC 2371  
Qy 2210 TCTGAGAAATCTCATGAGCCCAACCTTAACCTTAAGATCTTCACTTGGAGGACCAACC 2269  
Db 2372 TCTGAGAAATCTCATGAGCCCAACCTTAACCTTAAGATCTTCACTTGGAGGACCAACC 2431  
Qy 2270 TGAAGAAAGGATGAAGATGCGGTGAAGCTTAAACACCCAAATGTTGTTGG 2339  
Db 2432 TGAAGAAAGGATGAAGATGCGGTGAAGCTTAAACACCCAAATGTTGTTGG 2491  
Qy 2330 AGTCTTTGAGGCTGATTTGCTGTGATTTGAACCAATGCTGTTACTTGAAGATCTCCCAA 2389  
Db 2492 AGTCTTTGAGGCTGATTTGCTGTGATTTGAACCAATGCTGTTACTTGAAGATCTCCCAA 2551  
Qy 2390 TCTTTAGACCTCCCCAGCTGAAATCTCTGAGCTGCGGAGAGAA CAAAGTGA CAGACC 2449  
Db 2552 TCTTTAGACCTCCCCAGCTGAAATCTCTGAGCTGCGGAGAGAA CAAAGTGA CAGACC 2611  
Qy 2450 AGGAGTAATGCTCTGAGTATGCTTGAAGATCTCCAGTGGCGCTGAGAAAGCTGA 2509  
Db 2612 AGGAGTAATGCTCTGAGTATGCTTGAAGATCTCCAGTGGCGCTGAGAAAGCTGA 2671  
Qy 2510 TACTGAGAGACTGTGAGCTACACAGCCAGGATTCAGAGATCTTGACCTCGTCA 2569  
Db 2672 TACTGAGAGACTGTGAGCTACACAGCCAGGATTCAGAGATCTTGACCTCGTCA 2731  
Qy 2570 GCAACCGGAGCTTGA CACACTGTGCTATTCACAA CAGCTGGGGAAGAAAGTGTAA 2629  
Db 2732 GCAACCGGAGCTTGA CACACTGTGCTATTCACAA CAGCTGGGGAAGAAAGTGTAA 2791  
Qy 2630 ATCTACTGTGTGATTCATGAGGCTTCCCACTGTACTGCAAGAGCTGATGCTGAATC 2689  
Db 2792 ATCTACTGTGTGATTCATGAGGCTTCCCACTGTACTGCAAGAGCTGATGCTGAATC 2851  
Qy 2690 AGTGCACCTTGA CACCGCTGTGTGTTTCTTGTGACTTGTGCGCTTATGGTTAACTCAT 2749  
Db 2852 AGTGCACCTTGA CACCGCTGTGTGTTTCTTGTGACTTGTGCGCTTATGGTTAACTCAT 2911  
Qy 2750 GGTGAGCAGCACTGAGCTTATGATGAACCTGTGTGAAGCAATGCGTGAAGCTTCTGT 2809  
Db 2912 GGTGAGCAGCACTGAGCTTATGATGAACCTGTGTGAAGCAATGCGTGAAGCTTCTGT 2971  
Qy 2810 GCGAGCTATGAGAGAACATCTTGTATCTCCAGAGACTGAGATTTGTAAGTGTATC 2869  
Db 2972 GCGAGCTATGAGAGAACATCTTGTATCTCCAGAGACTGAGATTTGTAAGTGTATC 3031  
Qy 2870 TCAACCGCGGTGCTGTGAGAGTCTGTCTGTGTATCTGAGAGAGACCACTTGAAGA 2929  
Db 3032 TCAACCGCGGTGCTGTGAGAGTCTGTCTGTGTATCTGAGAGAGACCACTTGAAGA 3091  
Qy 2930 GCTGTGATCTCAACGACAAATGCCCTGGGTGACGGTGGGTTGCTGCGCTGTGCGAGGAC 2989  
Db 3092 GCTGTGATCTCAACGACAAATGCCCTGGGTGACGGTGGGTTGCTGCGCTGTGCGAGGAC 3151

Oy	2990	TGAAGCAAAAGAACGCTTCTCGACGAACTACCGGTTGAAGGCATGTGGACTGACTTCTG	3049
Db	3152	TGAACCAAAAGAACGCTTCTCGACGAACTACCGGTTGAAGGCATGTGGACTGACTTCTG	3211
Oy	3050	ATTGCTGTGAGGCACTCTCTTGGCCCTTTCCTGCAACCGGCATCTGACCAAGTCTTAAAC	3109
Db	3212	ATTGCTGTGAGGCACTCTCTTGGCCCTTTCCTGCAACCGGCATCTGACCAAGTCTTAAAC	3271
Oy	3110	TGTGTGACAATTACTTCAGTCCCAAGGAATGATGAACTGTGTGGCCCTTTGCCTGTC	3169
Db	3272	TGTGTGACAATTACTTCAGTCCCAAGGAATGATGAACTGTGTGTGGCCCTTTGCCTGTC	3331
Oy	3170	CCAGCTCTTAACCTTACAGATTAATTGGGCTGTGGAAATGCGACGTACCCCTGTGCAATTAAGGA	3229
Db	3332	CCAGCTCTTAACCTTACAGATTAATTGGGCTGTGGAAATGCGACGTACCCCTGTGCAATTAAGGA	3391
Oy	3230	AGCTGCTGTGAGGAAGTGCAGCTACCTCAAGCCCCGAGTGTAAATTGAACGTAGTTGGCAATT	3289
Db	3392	AGCTGCTGTGAGGAAGTGCAGCTACCTCAAGCCCCGAGTGTAAATTGAACGTAGTTGGCAATT	3451
Oy	3290	CTTTTGATGAAGATGACCG	3308
Db	3452	CTTTTGATGAAGATGACCG	3470

```

RESULT 5
US-10-216-645-3
Sequence 3, Application US/10216645
Publication No. US20030125282A1
GENERAL INFORMATION:
APPLICANT: WEISS, BERTRAM
APPLICANT: LESSL, MONIKA
APPLICANT: PETERS-KOTTIG, MICHAELE
APPLICANT: BECKMANN, GEORG
TITLE OF INVENTION: HUMAN MATRIX PROTEINS
FILE REFERENCE: SCH-1910
CURRENT APPLICATION NUMBER: US/10/216,645
PRIORITY FILING DATE: 2003-01-21
PRIOR APPLICATION NUMBER: DE 101 39 874.3 411
PRIORITY FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 3830
TYPE: DNA
ORGANISM: Homo sapiens
US-10-216-645-3

```

Query Match	70.9%	Score 2662	DB 15	Length 3830
Best Local Similarity	99.9%	Pred. No. 0		
Matches 2962	Conservative	0	Mismatches 2	Indels 0
Qy	AGAAATTTCCACAGCTATGGAAACAAGAGTCCACAGACGACAGACAGAAACAAG	404		
Db	AGAAATTTCCACAGCTATGGAAACAAGAGTCCACAGACGACAGACAGAAACAAG	509		
Qy	ACATGAGGTGACACATGCGACTCAAGAGTACCGTGTACCAATTGCTGAGGAGGA	464		
Db	ACATGAGGTGACACATGCGACTCAAGAGTACCGTGTACCAATTGCTGAGGAGGA	569		
Qy	GGATGTACGTGCTAGCTTTGAAAACACTGCTGCTGACTGCGCCGGAATGCAACGTTGGC	524		
Db	GGATGTACGTGCTAGCTTTGAAAACACTGCTGCTGACTGCGCCGGAATGCAACGTTGGC	629		
Qy	TGATGCTTTTGAATTACAGCCGGTGGGGCTTCCGGCTCGCAACGTTGTTCTGACCGAAA	584		
Db	TGATGCTTTTGAATTACAGCCGGTGGGGCTTCCGGCTCGCAACGTTGTTCTGACCGAAA	689		
Qy	GTCAAGAAATTGGGAAATGCGCTCTAGCCAGAGAAATCGTGTGCTGCGGCCCAACGTGG	644		
Db	GTCAAGAAATTGGGAAATGCGCTCTAGCCAGAGAAATCGTGTGCTGCGGCCCAACGTGG	749		

OY		645	ACTCTACAGGGAAAGTTCTCTCAAGTCTTCCTCCCGTTAAGAAGATGACAGGAA	704
Db		750	ACTTACCAAGGAAGTTCCTTACGTCTTCTCTCCCGTTAAGAAGATGACAGGAA	809
OY		705	GAAAGAGAGCAGTGTCACAAGTTTCATCTCCAGGAGTGGCCAGACTCCAGCTCCGGT	764
Db		810	GAAAGAGAGCAGTGTCACAAGTTTCATCTCCAGGAGAGTGGCCAGACTCCAGCTCCGGT	869
OY		765	GACGAGATCATGTCCGCACCAAAAAGGCTGTTGTTCAATGACGGTTTCATGACCT	824
Db		870	GACGAGATCATGTCCGCACCAAAAAGGCTGTTGTTCAATGACGGTTTCATGACCT	929
OY		825	GGGCTCTGTCCCAACATGACCAAACCTCTGCCAAAACCTGGGGCTGAAGCACCTCC	884
Db		930	GGGCTCTGTCCCAACATGACCAAACCTCTGCCAAAACCTGGGGCTGAAGCACCTCC	989
OY		885	GTTCACTCTCATACGACAGTCTGTGAGGAAGTCTGTCTCCCTGAGTCTTTCCTATCGT	944
Db		990	GTTCACTCTCATACGACAGTCTGTGAGGAAGTCTGTCTCCCTGAGTCTTTCCTATCGT	1044
OY		945	CACCGTCAGAGACGTGGGCACAGAGAACCTCAAGTCAAGAGTCTGTCTCCCGTTACCT	1000
Db		1050	CACCGTCAGAGACGTGGGCACAGAGAACCTCAAGTCAAGAGTCTGTCTCCCGTTACCT	1100
OY		1005	GTTAGTTAGAGGAATCTCCGGGGAAACAAGATCCAATTGCTCTTGAAGCGGGATTGG	1066
Db		1110	GTTAGTTAGAGGAATCTCCGGGGAAACAAGATCCAATTGCTCTTGAAGCGGGATTGG	1166
OY		1065	TGAGCATCAGAAACACAAGGGTTGCGGCGATCATGAAACAACGTTGAGCTGTGACCA	1122
Db		1170	TGAGCATCAGAAACACAAGGGTTGCGGCGATCATGAAACAACGTTGAGCTGTGACCA	1222
OY		1125	GTGCAGAGTGC CGCGCGTGGGCTCTCTCATCTGCGTGGCCCTGACGTGCAGAGCTGGT	1184
Db		1230	GTGCAGAGTGC CGCGCGTGGGCTCTCTCATCTGCGTGGCCCTGACGTGCAGAGCTGGT	1289
OY		1185	GGGGAGAGCGTCCGCCCTTCAACCAAACGCTCACAGGCTGCACGCCGCTTTGTGTT	1244
Db		1290	GGGGAGAGCGTCCGCCCTTCAACCAAACGCTCACAGGCTGCACGCCGCTTTGTGTT	1344
OY		1245	TCATAGCTCAACCCCTGAGGCGTGGGTCGGGCGTGTCAATCTGAGGAGAAAGATTGT	1300
Db		1350	TCATAGCTCAACCCCTGAGGCGTGGGTCGGGCGTGTCAATCTGAGGAGAAAGATTGT	1400
OY		1305	CCTGAAGCGCTTTCGCGTAGTGTGTGAGAGGAGTGTGGAATAAGAAAGTCAGTGTTTGA	1364
Db		1410	CCTGAAGCGCTTTCGCGTAGTGTGTGAGAGGAGTGTGGAATAAGAAAGTCAGTGTTTGA	1464
OY		1365	TGCTGACGACCTCATGTTCAAAGACTGGGGAAGTCTAGACCTCGTGCTCTGTTCACAT	1422
Db		1470	TGCTGACGACCTCATGTTCAAAGACTGGGGAAGTCTAGACCTCGTGCTCTGTTCACAT	1522
OY		1425	GAACTATCTTTCCTCCAGCACCCACTGTGAAGAGTACTAACCTTTCCTCCACTCAAGCT	1488
Db		1530	GAACTATCTTTCCTCCAGCACCCACTGTGAAGAGTACTAACCTTTCCTCCACTCAAGCT	1588
OY		1485	CCAGGACTTCTGTGCGCGCTTGTACTACGTGTAGAAGGCTGTGAAATTCAGACGACTCT	1544
Db		1590	CCAGGACTTCTGTGCGCGCTTGTACTACGTGTAGAAGGCTGTGAAATTCAGACGACTCT	1644
OY		1545	CTGCCCTCTGTACGTTGAAAGACAAGAGGTCCATGAGGCTTAAACAGGACGCTTCCA	1604
Db		1650	CTGCCCTCTGTACGTTGAAAGACAAGAGGTCCATGAGGCTTAAACAGGACGCTTCCA	1709
OY		1605	TATCCACTCTGCTTTTGAATGAAGCGTTTTTGTGTGCTCTGTGAGCCAAACGTTAAAGAG	1666
Db		1710	TATCCACTCTGCTTTTGAATGAAGCGTTTTTGTGTGCTCTGTGAGCCAAACGTTAAAGAG	1766
OY		1665	GCACTGGAAGGCTCTGCTGGGCTGTGCCGTTCCCTGGGGGTGAGACGAAAGCTTCTGCA	1722
Db		1770	GCACTGGAAGGCTCTGCTGGGCTGTGCCGTTCCCTGGGGGTGAGACGAAAGCTTCTGCA	1822
OY		1725	CTGGGCTCTCTGTGTGGGTGACAGACCTTAATGACCACCCACAGAGACACCTTGACGC	1784



```

CURRENT APPLICATION NUMBER: US/10/092,900A
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: US$N 60/274,322
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: US$N 60/283,675
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US$N 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: US$N 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: US$N 60/274,191
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: US$N 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US$N 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US$N 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US$N 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: US$N 60/287,424
PRIOR FILING DATE: 2001-04-10
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 768
SEQ ID NO 347
LENGTH: 3226
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (69)..(3168)
US-10-092-900A-347

```

Query Match	69.4%	Score 2801	DB 18	Length 3226
Best Local Similarity	99.9%	Pred. No. 0		
Matches 2501, Conservative	0	Mismatches 2	Indels 0	Gaps 0

Qy	406	CATGGAGGGTGAACATGGGACCTACAAAGACTCAGTATACCAAAATTCGTGAGGAG	465
Db	249	CATGAGGGTGAACATGGGACTACAAAGACTCAGTATGACCAAAATTCGTGAGGAG	308
Qy	466	GATGACGCTCGTAGTTTGTAAACAACCTGCTGCTGACTGGCCGGGAAATGCAAAAGTTGGCT	525
Db	309	GATGACGCTCGTAGTTTGTAAACAACCTGCTGCTGACTGGCCGGGAAATGCAAAAGTTGGCT	368
Qy	526	GGTGCTTTTGATTTCAGACCGGATGGGGCTTCCGGCTTCGACCGGTGTTCTGCACGGAAAG	585
Db	369	GGTCTTTTGATTTCAGACCGGATGGGGCTTCCGGCTTCGACCGGTGTTCTGCACGGAAAG	428
Qy	586	TCAGGAATTGGGAAATCGGCTTCAGCCAGAGGATCGTCTGTGCTGGCGCGAAGTGA	645
Db	429	TCAGGAATTGGGAAATCGGCTTCAGCCAGAGGATCGTCTGTGCTGGCGCGAAGTGA	488
Qy	646	CTTACCCAGGGAATGTTCTCTACGCTTCTTCTCCCGTTAGAGAGATGCACGGGAAG	705
Db	489	CTTACCCAGGGAATGTTCTCTACGCTTCTTCTCCCGTTAGAGAGATGCACGGGAAG	548
Qy	706	AAGGAGACAGTGTCAACAGATTCAATCTCCAGGAGTGGCCAGACTCCACGCTCCGGTG	765
Db	549	AAGGAGACAGTGTCAACAGATTCAATCTCCAGGAGTGGCCAGACTCCACGCTCCGGTG	608
Qy	766	ACGAGATCATGTCCCGACCAAGAAAGGCTGTTGTCATCATTTGACCGTTTCATGTACCTG	825
Db	609	ACGAGATCATGTCCCGACCAAGAAAGGCTGTTGTCATCATTTGACCGTTTCATGTACCTG	668
Qy	826	GGCTTGTTCCTCAACATGACACAAAGCTTCGCAAAAGACTGGGCTGAGAGCAGGCTCCG	885
Db	669	GGCTTGTTCCTCAACATGACACAAAGCTTCGCAAAAGACTGGGCTGAGAGCAGGCTCCG	728
Qy	886	TTTACCCCTCATACGACAGTCTGCTGAGGAAGGTCCTGCTCCCTGATGTCCTTCTGATTCGTC	945
Db	729	TTTACCCCTCATACGACAGTCTGCTGAGGAAGGTCCTGCTCCCTGATGTCCTTCTGATTCGTC	788

QY	946	ACCGGACAGACGTCGGGACACAGAGAAGTCGAATCAGAGAGTCGTCTCCCCGTTAACCTG	1005
Db	789	ACCGTCAGAGACGTGGGACACAGAGAAGTCGAATCAGAGAGTCGTCTCCCCGTTAACCTG	848
QY	1006	TTAGTTAAGGAATCTCCGGGGAAACAAAGAAATCCACTTGTCTCTTGACGCGGGATTGGT	1065
Db	849	TTAGTTAAGGAATCTCCGGGGAAACAAAGAAATCCACTTGTCTCTTGACGCGGGATTGGT	908
QY	1066	GAGCATCAGAAAGACAAAGGGTTGCGTCGATCAATGAACAACTGAGCTGTGTCAGCAG	1125
Db	909	GAGCATCAGAAAGACAAAGGGTTGCGTCGATCAATGAACAACTGAGCTGTGTCAGCAG	968
QY	1126	TGCCAGGTGCCCCGCTGGGCTCTCTCATCTGCGTGGCCCTGACCTGACAGACGTGGT	1185
Db	969	TGCCAGGTGCCCCGCTGGGCTCTCTCATCTGCGTGGCCCTGACCTGACAGACGTGGT	1028
QY	1186	GGGGAGAGCGTCGCCCCCTTAAACCAACGTCACAGGCTGACAGCCGCTTTGTGTT	1245
Db	1029	GGGGAGAGCGTCGCCCCCTTAAACCAACGTCACAGGCTGACAGCCGCTTTGTGTT	1088
QY	1246	CATCAGCTCACCCCTCGAGCGTGGTCCGCGCTGTCTCAATCTGAGAAAGATTGTG	1303
Db	1089	CATCAGCTCACCCCTCGAGCGTGGTCCGCGCTGTCTCAATCTGAGAAAGATTGTG	1145
QY	1306	CTGAAGCCCTTCTGCGTATATGCTGTGAGGAGATGTGGAAATAGAACTACGTGTTGAT	1365
Db	1149	CTGAAGCCCTTCTGCGTATATGCTGTGAGGAGATGTGGAAATAGAACTACGTGTTGAT	1208
QY	1366	GGTGACGACCTCATAGGTTCAAGGAACGTGGGGAGTGAAGCTCGGAGCTGTGTTCAATG	1422
Db	1209	GGTGACGACCTCATAGGTTCAAGGAACGTGGGGAGTGAAGCTCGGAGCTGTGTTCAATG	1266
QY	1426	AACATCTCTTCCGAGACAGCCACTGTGAGGAGTACTACACTTCTTCACTCAGTCTC	1485
Db	1269	AACATCTCTTCCGAGACAGCCACTGTGAGGAGTACTACACTTCTTCACTCAGTCTC	1328
QY	1486	CAGGACTTCTGTGCGGCTTGTACTACGTGTTAAGAGGCCCTGGAATCAGCCAGCTTC	1545
Db	1329	CAGGACTTCTGTGCGGCTTGTACTACGTGTTAAGAGGCCCTGGAATCAGCCAGCTTC	1388
QY	1546	TGCCCTCGTACGTTGAGAAACAAAGAGTCCATGAGCGTTAAACAGGACGGCTTCCAT	1605
Db	1389	TGCCCTCGTACGTTGAGAAACAAAGAGTCCATGAGCGTTAAACAGGACGGCTTCCAT	1448
QY	1606	ATCCACTCGCTTGGAGTAAGACGTTTTCTGTTGAGCCTCGAGACGGAAGCGTAAGAGG	1665
Db	1449	ATCCACTCGCTTGGAGTAAGACGTTTTCTGTTGAGCCTCGAGACGGAAGCGTAAGAGG	1508
QY	1666	CCACTGAGAGTCTGTGAGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAACCTTCTGCAC	1725
Db	1509	CCACTGAGAGTCTGTGAGGCTGTCCCGTTCCCTGGGGGTGAAGCAGAACCTTCTGCAC	1568
QY	1726	TGGGCTCTCTGTGTTGGGTCAAGACCTTAATGCCACCACTCGAGAGACCTTGAAGCC	1785
Db	1569	TGGGCTCTCTGTGTTGGGTCAAGACCTTAATGCCACCACTCGAGAGACCTTGAAGCC	1628
QY	1786	TTCCACTGTCTTTTGCAGACTCAAGACAAAGATTGTTGCTTGCGATTAAACAGCTTC	1845
Db	1629	TTCCACTGTCTTTTGCAGACTCAAGACAAAGATTGTTGCTTGCGATTAAACAGCTTC	1688
QY	1846	CAAGAAGTGTGGCTTCGGATTAAACCAAGACCTGGAATTGATAGACTTCTCTGTGCTC	1905
Db	1689	CAAGAAGTGTGGCTTCGGATTAAACCAAGACCTGGAATTGATAGACTTCTCTGTGCTC	1748
QY	1906	CAGCACTGTCCGTATTTTGGCGGAAATTCGGGTGATATCAAAAGGATCTTCCCAAGAT	1965
Db	1749	CAGCACTGTCCGTATTTTGGCGGAAATTCGGGTGATATCAAAAGGATCTTCCCAAGAT	1808
QY	1966	GAGTCGCGTGAAGGATGCTGTGAGTCCCTCTATAGATGCGGGATTAAAGCCCTCATTTGAG	2025
Db	1809	GAGTCGCGTGAAGGATGCTGTGAGTCCCTCTATAGATGCGGGATTAAAGCCCTCATTTGAG	1868
QY	2026	GAGCAGTGTGGAGATTTCTGTCTCATCTTGGACCAACCACTGCGGACCTGAC	2085





1046 AGATCATGTCGCCGACCAAGAAAGGCTGTTCATCATTTGACGGTTTCATGACCTGGGCT 1105  
QY 830 CTGTCTCAACAATGACACAAAGCTTCGAAAGACTGGGCTGAGAGACGCTCCGTTCA 889  
Db 1106 CTGTCTCAACAATGACACAAAGCTTCGAAAGACTGGGCTGAGAGACGCTCCGTTCA 1165  
QY 890 CCTCATAGGAGTCTGTCTGAGAGAGTCTGTCTCCATAGTCTTCTGATCTGTCACCG 949  
Db 1166 CCTCATAGGAGTCTGTCTGAGAGAGTCTGTCTCCATAGTCTTCTGATCTGTCACCG 1225  
QY 950 TCAGAGACGTGGGACACAGAGAGCTCAAGTCAAGAGTCTGTCTCCGTTACTGTTAG 1009  
Db 1226 TCAGAGACGTGGGACACAGAGAGCTCAAGAGTCTGTCTCCGTTACTGTTAG 1285  
QY 1010 TTAGAGAAATCTCCGGGAAACAAAGAAATCACTTGTCTCTTGAAGCGGGAATGTTGAGC 1069  
Db 1286 TTAGAGAAATCTCCGGGAAACAAAGAAATCACTTGTCTCTTGAAGCGGGAATGTTGAGC 1345  
QY 1070 ATCAGAGACACAGAGGTTGGTGGCATATGAAACAACGTGAGCTGCTGACACAGTGC 1129  
Db 1346 ATCAGAGACACAGAGGTTGGTGGCATATGAAACAACGTGAGCTGCTGACACAGTGC 1405  
QY 1130 AGGTGCGCGCGTGGGCTCTCATCTGAGTGGCCCTGACGCTGACAGAGTGTGGGAG 1189  
Db 1406 AGGTGCGCGCGTGGGCTCTCATCTGAGTGGCCCTGACGCTGACAGAGTGTGGGAG 1465  
QY 1190 AGAGGCTGCGCCCTTCAACAACGCTGACAGGCTGACGCGCTTGTGTTCATC 1249  
Db 1466 AGAGGCTGCGCCCTTCAACAACGCTGACAGGCTGACGCGCTTGTGTTCATC 1525  
QY 1250 AGCTCAACCTCTGAGGCGTGTCTCGGCGCTGTCTCATTTGAGAGAAAGTTGCTGCA 1309  
Db 1526 AGCTCAACCTCTGAGGCGTGTCTCGGCGCTGTCTCATTTGAGAGAAAGTTGCTGCA 1585  
QY 1310 AGCGCTTGGCGGTATGCTGTGAGAGAGTGGAAATGAGAAATGAGTTGATGGG 1369  
Db 1586 AGCGCTTGGCGGTATGCTGTGAGAGAGTGGAAATGAGAAATGAGTTGATGGG 1645  
QY 1370 AGACCTCATGTTCAAGAGACTCGGAGAGTGAAGTCCGCTCTGTTTCAATGACA 1429  
Db 1646 AGACCTCATGTTCAAGAGACTCGGAGAGTGAAGTCCGCTCTGTTTCAATGACA 1705  
QY 1430 TCTCTTCTCCAGACAGCCACTGTGAGAGTACTACACTTCTTCAACTCACTCTCCAG 1489  
Db 1706 TCTCTTCTCCAGACAGCCACTGTGAGAGTACTACACTTCTTCAACTCACTCTCCAG 1765  
QY 1490 ACTTCGTGCGGCTTGTATCTACGTGTGAGAGGCTGGAATCGAGCCAGCTCTGCG 1549  
Db 1766 ACTTCGTGCGGCTTGTATCTACGTGTGAGAGGCTGGAATCGAGCCAGCTCTGCG 1825  
QY 1550 CTCTGTACGTTGAGAGACAAAGAGTCCATGAGCTTAAACAGGAGGCTTCCATATCC 1609  
Db 1826 CTCTGTACGTTGAGAGACAAAGAGTCCATGAGCTTAAACAGGAGGCTTCCATATCC 1885  
QY 1610 ACTGCTTTGAGTGAAGCGTTTCTGTTGGCTCTGTGAGCGAAGACGTTAAGAGGCCAC 1669  
Db 1886 ACTGCTTTGAGTGAAGCGTTTCTGTTGGCTCTGTGAGCGAAGACGTTAAGAGGCCAC 1945  
QY 1670 TGGAGGCTCTGAGGCTGTCCCGTTCCTGAGGGGTGAAGAGAAAGCTTCTGACCTGG 1729  
Db 1946 TGGAGGCTCTGAGGCTGTCCCGTTCCTGAGGGGTGAAGAGAAAGCTTCTGACCTGG 2005  
QY 1730 TCTCTCTGTTGGGTGAGAGCTTAATGACACCCAGAGAGACACCTGAGCGCTTCC 1789  
Db 2006 TCTCTCTGTTGGGTGAGAGCTTAATGACACCCAGAGAGACACCTGAGCGCTTCC 2065  
QY 1790 ACTGCTTTTGAAGACTCAAGACAAAGATTGTTGCTTGGCATTTAAACAGCTTCCAG 1849  
Db 2066 ACTGCTTTTGAAGACTCAAGACAAAGATTGTTGCTTGGCATTTAAACAGCTTCCAG 2125  
QY 1850 AAGTGTGCTTCCGATTAACGAGACCTGGAATTGATATGCACTTCTTCTGCGCTCCAGC 1909

Db 2126 AAGTGTGCTTCCGATTAACGAGACCTGGAATTGATATGCACTTCTTCTGCGCTCCAGC 2185  
QY 1910 ACTGTCGGTATTTGGGAAAAATTGGGGTGAATGTCAAGAGGATCTTCCAAAGATGAGT 1969  
Db 2186 ACTGTCGGTATTTGGGAAAAATTGGGGTGAATGTCAAGAGGATCTTCCAAAGATGAGT 2245  
QY 1970 CCGTGAAGGCAATGTCTGTGGTCCCTCTAATGAAAGCGGGAATGAAGCCCTCATTTGAGAGC 2029  
Db 2246 CCGTGAAGGCAATGTCTGTGGTCCCTCTAATGAAAGCGGGAATGAAGCCCTCATTTGAGAGC 2305  
QY 2030 AGTGGAAAGATTTCTGTCTCAATGTCTGACCCACCACTGAGGAGGAGCTGACCTGG 2089  
Db 2306 AGTGGAAAGATTTCTGTCTCAATGTCTGACCCACCACTGAGGAGGAGCTGACCTGG 2365  
QY 2090 GCAGCAGATCTCTGACAGAGCGGCTCATGAAGACCTGTGTGCCAAGCTGAGGATCCCA 2149  
Db 2366 GCAGCAGATCTCTGACAGAGCGGCTCATGAAGACCTGTGTGCCAAGCTGAGGATCCCA 2425  
QY 2150 CCTGCAAGATACAGACCTGTGATTTGAATATGACAGATTAACCTGTGTGACGACAC 2209  
Db 2426 CCTGCAAGATACAGACCTGTGATTTGAATATGACAGATTAACCTGTGTGACGACAC 2485  
QY 2210 TCTGGAAGATGTCATGAGCACAACGTAACCTAAGTCCCTCAACTTGGGAGGACACCCAC 2269  
Db 2486 TCTGGAAGATGTCATGAGCACAACGTAACCTAAGTCCCTCAACTTGGGAGGACACCCAC 2545  
QY 2270 TGAAGAAAGAGATGAAGATGAGTGGTGAAGCTTAAACACCCAAATGTTTGTGG 2329  
Db 2546 TGAAGAAAGAGATGAAGATGAGTGGTGAAGCTTAAACACCCAAATGTTTGTGG 2605  
QY 2330 AGTCTTTAGGCTGATGTTGCTGTGATTTGACCCATGCTGTACTGAAAGATCTCCCAA 2389  
Db 2606 AGTCTTTAGGCTGATGTTGCTGTGATTTGACCCATGCTGTACTGAAAGATCTCCCAA 2665  
QY 2390 TCTTGAAGACCTCCCGGACGCTGAATCTGAGCTGGGACGAGAAACAAAGGTGACAGAC 2449  
Db 2666 TCTTGAAGACCTCCCGGACGCTGAATCTGAGCTGGGACGAGAAACAAAGGTGACAGAC 2725  
QY 2450 AAGGAGTAATGCTCTCAATGATGCTTGAAGTCTCCAGTGGCGCTGACAGAGCTGA 2509  
Db 2726 AAGGAGTAATGCTCTCAATGATGCTTGAAGTCTCCAGTGGCGCTGACAGAGCTGA 2785  
QY 2510 TACTGAGAGACTGTGAGCTACACAGCAGGCTGACAGAGTGTGGCTCAAGCTCTGCA 2569  
Db 2786 TACTGAGAGACTGTGAGCTACACAGCAGGCTGACAGAGTGTGGCTCAAGCTCTGCA 2845  
QY 2570 GCAACCGGAGCTTGAACACCTGTGCTATCCAAACACGCTGGGGAACGAGGTGTA 2629  
Db 2846 GCAACCGGAGCTTGAACACCTGTGCTATCCAAACACGCTGGGGAACGAGGTGTA 2905  
QY 2630 ATCTACTGTGTGATCCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATGCTGATC 2689  
Db 2906 ATCTACTGTGTGATCCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATGCTGATC 2965  
QY 2690 AGTGCACCTGGAACAGGCTGTGTGTTCTTGTGCACTTGCCTTATGGGTAACTCAT 2749  
Db 2966 AGTGCACCTGGAACAGGCTGTGTGTTCTTGTGCACTTGCCTTATGGGTAACTCAT 3025  
QY 2750 GGTGAGGACCTGAGCTTATGACATGAACCTGTGGAAGAAATGGGTGAAGCTTCTGT 2809  
Db 3026 GGTGAGGACCTGAGCTTATGACATGAACCTGTGGAAGAAATGGGTGAAGCTTCTGT 3085  
QY 2810 GCGAGGTATGAGAGAACATCTGTGATCTCCAGAGACCTGAGTGGTAAAGTGTCAATC 2869  
Db 3086 GCGAGGTATGAGAGAACATCTGTGATCTCCAGAGACCTGAGTGGTAAAGTGTCAATC 3145  
QY 2870 TCACCGCGCGTGTGTGAGAGTGTCTGTGTGATCTGAGAGAGACACCTGAAGA 2929  
Db 3146 TCACCGCGCGTGTGTGAGAGTGTCTGTGTGATCTGAGAGAGACACCTGAAGA 3205  
QY 2930 GCTGGAATCTGACGAGACAAATGCTGTGGGTGAAGGTTGGCTGTGGAGAGGAC 2989  
Db 3206 GCTGGAATCTGACGAGACAAATGCTGTGGGTGAAGGTTGGCTGTGGAGAGGAC 3265



1550 CTCTGTACCTTGAGAGCAAGAGGTCATGAGCTTAAACAGGAGCTTCCATATCC 1609  
1826 CTCTGTACCTTGAGAGCAAGAGGTCATGAGCTTAAACAGGAGCTTCCATATCC 1885  
1610 ACTGCTTTGAGAGGAGGCTTTCTGTTTGGCTTGTGAGCGAAGAGCTTAAGAGGCGAC 1669  
1886 ACTGCTTTGAGAGGAGGCTTTCTGTTTGGCTTGTGAGCGAAGAGCTTAAGAGGCGAC 1945  
1670 TGGAGGCTCTGCTGGGCTGTCCCGTTCCTTGGGGGTGAAGCAAGAGCTTCTGCACTGG 1729  
1946 TGGAGGCTCTGCTGGGCTGTCCCGTTCCTTGGGGGTGAAGCAAGAGCTTCTGCACTGG 2005  
1730 TCTCTGTGTGGGTGAGAGGCTTAATGACCAACCCAGAGAGCAACCTTGAAGGCTTCC 1789  
2006 TCTCTGTGTGGGTGAGAGGCTTAATGACCAACCCAGAGAGCAACCTTGAAGGCTTCC 2065  
1790 ACTGCTTTGAGAGCTCAAGCAAGAGGCTTGTGCTTGGAGCTTAAACAGCTTCAAG 1849  
2066 ACTGCTTTGAGAGCTCAAGCAAGAGGCTTGTGCTTGGAGCTTAAACAGCTTCAAG 2125  
1850 AAGTGTGCTTCCGATTAAACGAACCTGAGCTTGAATGACATCTTCTGCTTCCAGC 1909  
2126 AAGTGTGCTTCCGATTAAACGAACCTGAGCTTGAATGACATCTTCTGCTTCCAGC 2185  
1910 ACTGTCCGATTGTGGGAGAAATTCGGGTGGATGTCAAGGAGATCTTCCCAAGAGTAGT 1969  
2186 ACTGTCCGATTGTGGGAGAAATTCGGGTGGATGTCAAGGAGATCTTCCCAAGAGTAGT 2245  
1970 CCGCTGAGGAGATGCTCTGTGTCTCCCTTAATGAGTGGGAGTAAGACCTCATTTGAGAGC 2029  
2246 CCGCTGAGGAGATGCTCTGTGTCTCCCTTAATGAGTGGGAGTAAGACCTCATTTGAGAGC 2305  
2030 AAGTGGAGATTTCTGTCTCAATGCTTGGAGCCCAACCTGCGGAGCTGAGACTGG 2089  
2306 AAGTGGAGATTTCTGTCTCAATGCTTGGAGCCCAACCTGCGGAGCTGAGACTGG 2365  
2090 GAGAGGAGATCTCTGAGAGGAGGAGGAGAGAGCTGTGTGAGCAAGCTGAGAGCTTCCA 2149  
2366 GAGAGGAGATCTCTGAGAGGAGGAGGAGAGAGCTGTGTGAGCAAGCTGAGAGCTTCCA 2425  
2150 CCGTGAAGATACAGAGCCCTGATGTTAGAAATGACAGATTAACCTCTGTGTGAGAGC 2209  
2426 CCGTGAAGATACAGAGCCCTGATGTTAGAAATGACAGATTAACCTCTGTGTGAGAGC 2485  
2210 TCTGAGAGATGCTCATGAGCAACCTGATCTTAAGATCTCTCAACTTGGAGGAGCCAC 2269  
2486 TCTGAGAGATGCTCATGAGCAACCTGATCTTAAGATCTCTCAACTTGGAGGAGCCAC 2545  
2270 TGAAGGAGAGATGTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2329  
2546 TGAAGGAGAGATGTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2605  
2330 AGCTTTTGGAGGCTGATGCTGAGATTAACCATGAGCTGTTAACTGAGATCTCCCAA 2389  
2606 AGCTTTTGGAGGCTGATGCTGAGATTAACCATGAGCTGTTAACTGAGATCTCCCAA 2665  
2390 TCTTGAAGAGCTTCCAGAGCTGAAATCTGTAGGCTGAGAGGAGGAGGAGGAGGAGGAG 2449  
2666 TCTTGAAGAGCTTCCAGAGCTGAAATCTGTAGGCTGAGAGGAGGAGGAGGAGGAGGAG 2725  
2450 AGGAGGATATGCTCTCAAGTATGCTTGAAGAGTCTCCAGAGGAGGAGGAGGAGGAGGAG 2509  
2726 AGGAGGATATGCTCTCAAGTATGCTTGAAGAGTCTCCAGAGGAGGAGGAGGAGGAGGAG 2785  
2510 TACTGAGAGAGCTGTGGAGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2569  
2786 TACTGAGAGAGCTGTGGAGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2845  
2570 GCAACCGAGGCTTGAACAGCTGTGTCTTAACCAACAGCTTGGGAGGAGGAGGAGGAG 2629  
2846 GCAACCGAGGCTTGAACAGCTGTGTCTTAACCAACAGCTTGGGAGGAGGAGGAGGAG 2905  
2630 ATCTACTGTGTGATCCATGAGGCTTCCCACTGATAGTGTGAGAGGCTGATGCTGAATC 2689

2906 ATCTACTGTGTGATCCATGAGGCTTCCCACTGATAGTGTGAGAGGCTGATGCTGAATC 2965  
2690 AGTGCACCTTGAGAGAGGCTGTGTGTCTTGTGAGCTTGGCTTAATGGGTAATCTAT 2749  
2966 AGTGCACCTTGAGAGAGGCTGTGTGTCTTGTGAGCTTGGCTTAATGGGTAATCTAT 3025  
2750 GGTGAGGAGCTGAGGCTTGAAGAGGCTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2809  
3026 GGTGAGGAGCTGAGGCTTGAAGAGGCTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 3085  
2810 GCGAGGATGAGAGAGGCTTGTGATCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2869  
3086 GCGAGGATGAGAGAGGCTTGTGATCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3145  
2870 TCACCGCGGCTGTGTGAGAGGCTGTGTGTGATCTGAGAGGAGGAGGAGGAGGAGGAGGAG 2929  
3146 TCACCGCGGCTGTGTGAGAGGCTGTGTGTGATCTGAGAGGAGGAGGAGGAGGAGGAGGAG 3205  
2930 GCTGAGATCTCAAGGAGGAGGCTGTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2989  
3206 GCTGAGATCTCAAGGAGGAGGCTGTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3265  
2990 TGAAGCAAAAGAGAGGCTTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3049  
3266 TGAAGCAAAAGAGAGGCTTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3325  
3050 ATTGTGTGAGGAGGCTTCTCTTGGGCTTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3109  
3326 ATTGTGTGAGGAGGCTTCTCTTGGGCTTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3385  
3110 TGTGAGAGATTAATCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3169  
3386 TGTGAGAGATTAATCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3445  
3170 CCAAGCTTAACCTTCAAGATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3229  
3446 CCAAGCTTAACCTTCAAGATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3505  
3220 AGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3289  
3506 AGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3565  
3290 CTTTGTGATGAAGATGACCG 3308  
3566 CTTTGTGATGAAGATGACCG 3584

RESULT 9  
US-10-677-943-23  
Sequence 23, Application US/10677943  
Publication No. US20040072297A1  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America as  
APPLICANT: represented by the Secretary of the Department of Health and  
APPLICANT: Human Services  
APPLICANT: Nelson, Lawrence  
APPLICANT: Tong, Zhi-Bin  
TITLE OF INVENTION: Human Gene Critical to Fertility  
FILE REFERENCE: 4239-64790  
CURRENT APPLICATION NUMBER: US/10/677,943  
CURRENT FILING DATE: 2003-10-01  
PRIOR APPLICATION NUMBER: 60/241,510  
PRIOR FILING DATE: 2000-10-18  
PRIOR APPLICATION NUMBER: PCT/US02/09776  
PRIOR FILING DATE: 2002-03-29  
PRIOR APPLICATION NUMBER: PCT/US01/10981  
PRIOR FILING DATE: 2001-04-04  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 23  
LENGTH: 3900  
TYPE: DNA



2330 AGCTTTGAGGCTGATGCTGATGATGACCAATGCTGTTACCTGAAGATCTCCCAA 2389  
2606 AGCTTTGAGGCTGATGCTGATGATGACCAATGCTGTTACCTGAAGATCTCCCAA 2665  
2390 TCCCTTACGACCTCCCGACCTGAAATCTCTGAGCCTGACAGAAACAAGTGAACACC 2449  
2666 TCCCTTACGACCTCCCGACCTGAAATCTCTGAGCCTGACAGAAACAAGTGAACACC 2725  
2450 AGGAGATATGCTCTGATGATGCTGATGAGTCTCCAGAGTCCCGCTGCAAGAGTGA 2509  
2726 AGGAGATATGCTCTGATGATGCTGATGAGTCTCCAGAGTCCCGCTGCAAGAGTGA 2785  
2510 TACTGAGGAGCTGTGACATCAAGCAGGAGTTCAGAGTCCGAGCTGAGCCTGATCA 2569  
2786 TACTGAGGAGCTGTGACATCAAGCAGGAGTTCAGAGTCCGAGCTGAGCCTGATCA 2845  
2570 GCAACCGGAGCTTGAACAACCTGTGCTTATCCAAACAAGCCTGAGGAAAGAGTGA 2629  
2846 GCAACCGGAGCTTGAACAACCTGTGCTTATCCAAACAAGCCTGAGGAAAGAGTGA 2905  
2630 ATCTACTGTGTGATTCATGAGGCTTCCCACTGTATGTCGAGAGGCTGATGCTGATC 2689  
2906 ATCTACTGTGTGATTCATGAGGCTTCCCACTGTATGTCGAGAGGCTGATGCTGATC 2965  
2690 AGTGCACCTGGAACAGGCTGCTGCTGCTTCTTGTGCACTTGCGCTTATGAGTATCAT 2749  
2966 AGTGCACCTGGAACAGGCTGCTGCTGCTTCTTGTGCACTTGCGCTTATGAGTATCAT 3025  
2750 GGCCTGACGACCTGAGCCTTATGACATGAACTCTGTGAGAGACAATGAGCCTTCTGT 2809  
3026 GGCCTGACGACCTGAGCCTTATGACATGAACTCTGTGAGAGACAATGAGCCTTCTGT 3085  
2810 GCGAGGTCTATGAGAGAACCTATCTTGTATCTCCAGAGCCTGAGATTGTAAGTGTATC 2869  
3086 GCGAGGTCTATGAGAGAACCTATCTTGTATCTCCAGAGCCTGAGATTGTAAGTGTATC 3145  
2870 TCAACGCGCGCTGCTGATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2929  
3146 TCAACGCGCGCTGCTGATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3205  
2930 GGCCTGATCTCAAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2989  
3206 GGCCTGATCTCAAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3265  
2990 TGAAGCAAAAGAAAGAGTGTCTGAGAGAGCTGCGGCTTGAAGGCAATGAGTGAATCTTG 3049  
3266 TGAAGCAAAAGAAAGAGTGTCTGAGAGAGCTGCGGCTTGAAGGCAATGAGTGAATCTTG 3325  
3050 ATTGCTGAGAGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3109  
3326 ATTGCTGAGAGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3385  
3110 TGGTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3169  
3386 TGGTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3445  
3170 CCAAGCTTAACTTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3229  
3446 CCAAGCTTAACTTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3505  
3230 AGCTCTGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3289  
3506 AGCTCTGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3565  
3290 CTTTGTATGAGATGACCG 3308  
3566 CTTTGTATGAGATGACCG 3584

RESULT 10  
US-10-399-443-1  
; Sequence 1, Application US/10399443  
; Publication No. US20040028669A1

GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as Represented by the  
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
; APPLICANT: Health  
; APPLICANT: Nelson, Lawrence M.  
; APPLICANT: Tong, Zhi-Bin  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Zhi-Bin, Tong  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64785  
; CURRENT FILING DATE: 2003-04-16  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO: 1  
; LENGTH: 1157  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-399-443-1  
Query Match 27.3%; Score 1102; DB 17; Length 1157;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
757 GCTCCGAGAGAGAGATCATGCTCCGACCAAAAGGCTGTTTATCATGATGAGGTTTC 816  
4 GCTCCGAGAGAGAGATCATGCTCCGACCAAAAGGCTGTTTATCATGATGAGGTTTC 63  
817 GATGACCTGAGGCTCTGCTCTCAACATGACCAAAAGCTCTGCAAAAGCTGAGAG 876  
64 GATGACCTGAGGCTCTGCTCTCAACATGACCAAAAGCTCTGCAAAAGCTGAGAG 123  
877 CAGCTCCGTTCACTCTATACGAGTCTGCTGAGAGAGTCTGCTCTGAGTCTTC 936  
124 CAGCTCCGTTCACTCTATACGAGTCTGCTGAGAGAGTCTGCTCTGAGTCTTC 183  
937 CTGATGTCACCGTCAGAGAGTGGGACAGAGAGCTCAAGTCAAGAGTGTGCTCC 996  
184 CTGATGTCACCGTCAGAGAGTGGGACAGAGAGCTCAAGTCAAGAGTGTGCTCC 243  
997 GCTTACTGTTAGTTAGAGAACTCTCCGAGGAAACAAGATCACTTGTCTTGAAGGC 1056  
244 GCTTACTGTTAGTTAGAGAACTCTCCGAGGAAACAAGATCACTTGTCTTGAAGGC 303  
1057 GGGATGTTGATGATCAAGAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1116  
304 GGGATGTTGATGATCAAGAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 363  
1117 CTGACAGAGTCCAGAGTCCCGCGTGGGCTCTGATCTGCTGAGGCTCTGAGCTGAG 1176  
364 CTGACAGAGTCCAGAGTCCCGCGTGGGCTCTGATCTGCTGAGGCTCTGAGCTGAG 423  
1177 GAGCTGTGGGAGAGAGGCTGCGCCCTTCAACAAAGCTCAAGAGGCTGACGCGCT 1236  
424 GAGCTGTGGGAGAGAGGCTGCGCCCTTCAACAAAGCTCAAGAGGCTGACGCGCT 483  
1237 TTTGTGTTTATCAAGCTCACTCTGAGAGGCTGCTGAGGCTGCTGCTGCTGCTGCTGCTG 1296  
484 TTTGTGTTTATCAAGCTCACTCTGAGAGGCTGCTGAGGCTGCTGCTGCTGCTGCTGCTG 543  
1297 AGAGTGTCTGAGAGGCTTCTGCGATGAGCTGAGAGGAGTGTGGAATAGGAAGTCA 1356  
544 AGAGTGTCTGAGAGGCTTCTGCGATGAGCTGAGAGGAGTGTGGAATAGGAAGTCA 603  
1357 GTGTTTATGATGAGAGCTTATGAGTGTCAAGAGCTGAGAGTGTGAGTGTGCTGCTG 1416  
604 GTGTTTATGATGAGAGCTTATGAGTGTCAAGAGCTGAGAGTGTGAGTGTGCTGCTG 663  
1417 TTTCAATGAACATCTTCTTCCCAAGACGACATGTGAGAGATCTACATCTTCTTCCAC 1476

Db 664 TTTGACATGAACATCTTCTCCAGACAGCCACTGAGAGATCTAACACTTTCTTCAC 723  
Qy 1477 CTCAGCTTCAGAGACTTCTGTGCGCCCTTTGTACTAAGTTAGAGAGGCTTGAATTCGAG 1536  
Db 724 CTCAGCTTCAGAGACTTCTGTGCGCCCTTTGTACTAAGTTAGAGAGGCTTGAATTCGAG 783  
Qy 1537 CCAGCTCTCTGCGCTCTGTAGCTTGAAGAACAAAGAGGTCATGAGCTTAAACAGGCA 1596  
Db 784 CCAGCTCTCTGCGCTCTGTAGCTTGAAGAACAAAGAGGTCATGAGCTTAAACAGGCA 843  
Qy 1597 GGCTTCATATCCACTCGCTTTGATGAAGCGTTTCCTGTTGACCTGAGCGAAGC 1656  
Db 844 GGCTTCATATCCACTCGCTTTGATGAAGCGTTTCCTGTTGACCTGAGCGAAGC 903  
Qy 1657 GTAAAGAGGCCACTGAGAGTCTGTGCGCTTCCCGTTCCCTGAGGGGTGAAGCAGAG 1716  
Db 904 GTAAAGAGGCCACTGAGAGTCTGTGCGCTTCCCGTTCCCGTTGAGGGGTGAAGCAGAG 963  
Qy 1717 CTTCGCACTGAGGCTCTCTGTGTTGAGTCAAGACCTTAATGCCACCCAGAGACACC 1776  
Db 964 CTTCGCACTGAGGCTCTCTGTGTTGAGTCAAGACCTTAATGCCACCCAGAGACACC 1023  
Qy 1777 CTGAAGCCTTCACACTGTCTTTTTCAGACTCAAGACAAAGAGTTTTCCTGTCATTA 1836  
Db 1024 CTGAAGCCTTCACACTGTCTTTTTCAGACTCAAGACAAAGAGTTTTCCTGTCATTA 1083  
Qy 1837 AACACCTTCAGAGAGTGTGCTTCCGANTTACAGAAACCTGAGACTTGAATGACCTTCC 1896  
Db 1084 AACACCTTCAGAGAGTGTGCTTCCGANTTACAGAAACCTGAGACTTGAATGACCTTCC 1143  
Qy 1897 TTCTGCTTCAGC 1909  
Db 1144 TTCTGCTTCAGC 1156  
RESULT 11  
US-10-677-943-1  
; Sequence 1, Application US/10677943  
; Publication No. US20040072297A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as  
; APPLICANT: represented by the Secretary of the Department of Health and  
; APPLICANT: Human Services  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Tong, Zhi-Bin  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64790  
; CURRENT FILING DATE: 2003-10-01  
; PRIOR FILING DATE: 2003-10-01  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR FILING DATE: 2002-03-29  
; PRIOR FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; PRIOR FILING DATE: 2001-04-04  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 1157  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-677-943-1  
Query Match 27.3%; Score 1102; DB 18; Length 1157;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 757 GGTCCGGTGAAGGATGATGTCGCGACCAAGAGGCTGTTGATGATGAGCGTTTC 816  
Db 4 GTTCGGTGAAGGATGATGTCGCGACCAAGAGGCTGTTGATGATGAGCGTTTC 63  
Qy 817 GATGACCTGAGGCTGTGCTCTCAACAATGACAAAGCTCTGCAAGAGCTGGGCTGAGAG 876

Db 64 GATGACCTGAGGCTGTGCTCTCAACAATGACAAAGCTCTGCAAGAGCTGGGCTGAGAG 123  
Qy 877 CAGCCTCCGTTCAACCTTCATACGAGTCTGTAGAGAAAGTCTGCTCCCTGAATCTTC 936  
Db 124 CAGCCTCCGTTCAACCTTCATACGAGTCTGTAGAGAAAGTCTGCTCCCTGAATCTTC 183  
Qy 937 CTGATCGTCAACCTCAGAGAGTGGGCAAGAGAGGCTCAAGTCAAGAGTGTGTCTCC 996  
Db 184 CTGATCGTCAACCTCAGAGAGTGGGCAAGAGAGGCTCAAGTCAAGAGTGTGTCTCC 243  
Qy 997 CGTTACCTGTTAGTTAAGAAATCTCCGGGAAACAAGAAATCACTTGCTCTTGAAGGC 1056  
Db 244 CGTTACCTGTTAGTTAAGAAATCTCCGGGAAACAAGAAATCACTTGCTCTTGAAGGC 303  
Qy 1057 GGGATGGTGAAGCATCAAGAAACAAGAGGTTGGTGCATCAATGAACAACCGTGAAG 1116  
Db 304 GGGATGGTGAAGCATCAAGAAACAAGAGGTTGGTGCATCAATGAACAACCGTGAAG 363  
Qy 1117 CTGACACGATGCAAGTGCAGGCTCCGCGTGGGCTCTCATCTGCGTGCCTGACCTGAG 1176  
Db 364 CTGACACGATGCAAGTGCAGGCTCCGCGTGGGCTCTCATCTGCGTGCCTGACCTGAG 423  
Qy 1177 GACGTGGTGGGAGAGGCTGCCCTTCAACCAACGCTCAGAGGCTTGCAGCGGCT 1236  
Db 424 GACGTGGTGGGAGAGGCTGCCCTTCAACCAACGCTCAGAGGCTTGCAGCGGCT 483  
Qy 1237 TTTGTGTTTCAATCAAGCTCAACCCCTGAGAGGCTGGTCCGGGCTGCTCATCTGAGAG 1296  
Db 484 TTTGTGTTTCAATCAAGCTCAACCCCTGAGAGGCTGGTCCGGGCTGCTCATCTGAGAG 543  
Qy 1297 AGAGTTGCTGAAGAGCTTCTGCGGATGAGCTGAGAGGAGTGAATAGAGATCA 1356  
Db 544 AGAGTTGCTGAAGAGCTTCTGCGGATGAGCTGAGAGGAGTGAATAGAGATCA 603  
Qy 1357 GTGTTTGAATGATGAGAGCTTCATGATGTTCAAGAGCTGGGAGTCTGAGCTCTG 1416  
Db 604 GTGTTTGAATGATGAGAGCTTCATGATGTTCAAGAGCTGGGAGTCTGAGCTCTG 663  
Qy 1417 TTTTCAATGAACATCTTCTTCCAGACAGCACTGTAGAGAGTCTACACTTCTTCAC 1476  
Db 664 TTTTCAATGAACATCTTCTTCCAGACAGCACTGTAGAGAGTCTACACTTCTTCAC 723  
Qy 1477 CTCAGCTTCAGAGACTTGTGCGCGCTTGAAGTCAAGTGTAGAGGCTTGAATTCGAG 1536  
Db 724 CTCAGCTTCAGAGACTTGTGCGCGCTTGAAGTCAAGTGTAGAGGCTTGAATTCGAG 783  
Qy 1537 CCAGCTCTCTGCGCTCTGTAGAGTGAAGAACAAAGAGTTCATGAGCTTAAACAGGCA 1596  
Db 784 CCAGCTCTCTGCGCTCTGTAGAGTGAAGAACAAAGAGTTCATGAGCTTAAACAGGCA 843  
Qy 1597 GGCTTCATATCCACTCGCTTTGATGAAGCGTTTCCTGTTGACCTGAGCGAAGC 1656  
Db 844 GGCTTCATATCCACTCGCTTTGATGAAGCGTTTCCTGTTGACCTGAGCGAAGC 903  
Qy 1657 GTAAAGAGGCCACTGAGAGTCTGTGCGCTTCCCGTTCCCGTTGAGGGGTGAAGCAGAG 1716  
Db 904 GTAAAGAGGCCACTGAGAGTCTGTGCGCTTCCCGTTCCCGTTGAGGGGTGAAGCAGAG 963  
Qy 1717 CTTCGCACTGAGGCTCTCTGTGTTGAGTCAAGACCTTAATGCCACCCAGAGACACC 1776  
Db 964 CTTCGCACTGAGGCTCTCTGTGTTGAGTCAAGACCTTAATGCCACCCAGAGACACC 1023  
Qy 1777 CTGAAGCCTTCACACTGTCTTTTTCAGACTCAAGACAAAGAGTTTTCCTGTCATTA 1836  
Db 1024 CTGAAGCCTTCACACTGTCTTTTTCAGACTCAAGACAAAGAGTTTTCCTGTCATTA 1083  
Qy 1837 AACACCTTCAGAGAGTGTGCTTCCGANTTACAGAAACCTGAGACTTGAATGACCTTCC 1896  
Db 1084 AACACCTTCAGAGAGTGTGCTTCCGANTTACAGAAACCTGAGACTTGAATGACCTTCC 1143  
Qy 1897 TTCTGCTTCAGC 1909



```

Db          1144  TTCTGCTTCACG  1156

RESULT 12
US-10-399-443-3
; Sequence 3, Application US/10399443
; Publication No. US20040028669A1
GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as Represented by the
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute
; APPLICANT: Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
TITLE OF INVENTION: Human Gene Critical to Fertility
FILE REFERENCE: 4239-64785
CURRENT APPLICATION NUMBER: US/10/399,443
CURRENT FILING DATE: 2003-04-16
PRIOR APPLICATION NUMBER: 60/241,510
PRIOR FILING DATE: 2000-10-18
PRIOR APPLICATION NUMBER: PCT/US01/10981
PRIOR FILING DATE: 2001-04-04
NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1075
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-399-443-3

```

Query Match	Similarity	16.3%	Pred 657;	DB 17;	Length 1075;
Best Local	Similarity	99.7%	Pred No. 0;		
Matches	757;	Conservative	0;	Mismatches	2;
				Indels	0;
				Gaps	0;
Qy	2550	TCGCGCTCAGCCCTGTCAGCAACCGGAGCTTGACACACCTGTGCTTATCCAACACAG	2609		
Db	1	TCGTGCTCTCAGCCCTGTCAGCAACCGGAGCTTGACACACCTGTGCTTATCCAACACAG	60		
Qy	2610	CCTGGGGAAACGAAGGTGTAAATCTATCTGTGTGATCCATGTAGAGGCTTCCCACTGAAGCT	2669		
Db	61	CTTGGGGAAACGAAGGTGTAAATCTATCTGTGTGATCCATGTAGAGGCTTCCCACTGAAGCT	120		
Qy	2670	GCAGAGGCTGATGCTGAATCAAGTGCACACTGACACACGCTGGCTGTGTTTTCTTGCACT	2729		
Db	121	GCAGAGGCTGATGCTGAATCAAGTGCACACTGACACACGCTGGCTGTGTTTTCTTGCACT	180		
Qy	2730	TGCGCTTAATGGGTAACTGATGGCTGAGGCACTGAGCCTTGACATGAAACCTGTGGAAGA	2789		
Db	181	TGCGCTTAATGGGTAACTGATGGCTGAGGCACTGAGCCTTGACATGAAACCTGTGGAAGA	240		
Qy	2790	CAATGGCCTGAAGCTTCTGTGCGAGAGGTATGAGAGAACATCTTGTCAATCTCCAGAACCT	2849		
Db	241	CAATGGCCTGAAGCTTCTGTGCGAGAGGTATGAGAGAACATCTTGTCAATCTCCAGAACCT	300		
Qy	2850	GGAATGGTAAAGTGTCAATCTCACCGCCGCTGCTGTGAGAGTGTCTCTGTGTGATCTC	2909		
Db	301	GGAATGGTAAAGTGTCAATCTCACCGCCGCTGCTGTGAGAGTGTCTCTGTGTGATCTC	360		
Qy	2910	GAGAGCGACGACACCTGAAGAGGCTTGATCTCAACGGACAAATGCCCTGGGTGAACGGTGGGCT	2969		
Db	361	GAGAGCGACGACACCTGAAGAGGCTTGATCTCAACGGACAAATGCCCTGGGTGAACGGTGGGCT	420		
Qy	2970	TGCTGCGCTGTGCGAGGACCTGAGCAAAAGAACAGTGTCTTGACGAGACTCGGGTGA	3029		
Db	421	TGCTGCGCTGTGCGAGGACCTGAGCAAAAGAACAGTGTCTTGACGAGACTCGGGTGA	480		
Qy	3030	GGAATGTGGAATGATCTGATGTGTGAGGCACTCCTTGTGCGCTTCTCTGCAACG	3089		
Db	481	GGAATGTGGAATGATCTGATGTGTGAGGCACTCCTTGTGCGCTTCTCTGCAACG	540		
Qy	3090	GCATCTGACCAAGTCTAAACCTGTGTGACAAATTAATTCAAGTCCCAAGGAATGATGAAGCT	3149		

Db	541	GCATCTGACGAGTCTTAACCTGGTGCAGAAATTACTTCACTGCCAAGAATGATGAAGCT	600
Oy	3150	GTGTTGGGCGCTTTGCTGCTGCCAGTCTTAAGTAATTTGGGCTGTGGAAATGCA	3205
Db	601	GTGTTGGGCGCTTTGCTGCTGCCAGTCTTAACAGTAATTTGGGCTGTGGAAATGCA	660
Oy	3210	GTACCCGTGTGGAATTAAGGAGCTGTGGAGGAAGTGAAGCTACACAGCCCGAGTGT	3265
Db	661	GTACCCGTGTGGAATTAAGGAGCTGTGGAGGAAGTGAAGCTACACAGCCCGAGTGT	720
Oy	3270	AATTGACGGTAGTGGCAATCTTTTGATGAAGATGACCG	3308
Db	721	AATTGACGGTAGTGGCAATCTTTTGATGAAGATGACCG	759

Query Match	16.3%; Score 657; DB 18; Length 1075;
Best Local Similarity	99.7%; Pred. No. 0;
Matches	757; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
US-10-677-943-3	US-10-677-943-3
Sequence 3, Application US/10677943	
Publication No. US20040072297A1	
GENERAL INFORMATION:	
APPLICANT: The Government of the United States of America as	
APPLICANT: represented by the Secretary of the Department of Health and	
APPLICANT: Human Services	
APPLICANT: Nelson, Lawrence	
APPLICANT: Tong, Zhi-Bin	
TITLE OF INVENTION: Human Gene Critical to Fertility	
FILE REFERENCE: 4239-64790	
CURRENT APPLICATION NUMBER: US/10/677,943	
PRIOR FILING DATE: 2003-10-01	
PRIOR APPLICATION NUMBER: 60/241,510	
PRIOR FILING DATE: 2000-10-18	
PRIOR APPLICATION NUMBER: PCT/US02/09776	
PRIOR FILING DATE: 2002-03-29	
PRIOR APPLICATION NUMBER: PCT/US01/10981	
PRIOR FILING DATE: 2001-04-04	
NUMBER OF SEQ ID NOS: 42	
SOFTWARE: PatentIn version 3.2	
SEQ ID NO 3	
LENGTH: 1075	
TYPE: DNA	
ORGANISM: Homo sapiens	
US-10-677-943-3	
2550 TCTGACCTCAGCCCTCGTCAGCAACCGAGCTTGACACACCTGTGCTATCCACACAACG	2605
1 TCTGACCTCAGCCCTCGTCAGCAACCGAGCTTGACACACCTGTGCTATCCACACAACG	60
2610 CCTGGGAAACCAAGGTGTAAATCTAATCTGTGTGATCCATGAGGCTTCCCACTGATGCT	2665
61 CCTGGGAAACCAAGGTGTAAATCTAATCTGTGTGATCCATGAGGCTTCCCACTGATGCT	120
2670 GCAGAGGCTGATGCTGATCAGTCGACCTGACACCGCTGAGTGTGTTTCTTGCACT	2725
121 GCAGAGGCTGATGCTGATCAGTCGACCTGACACCGCTGAGTGTGTTTCTTGCACT	180
2730 TGCGCTTATGAGGTAACTCATGAGCTGACGACCTGAGCCTTGAATGAACCCCTGTGAGA	2785
181 TGCGCTTATGAGGTAACTCATGAGCTGACGACCTGAGCCTTGAATGAACCCCTGTGAGA	240
2790 CAATGGCTGTAAGCTTCTGTGCCAGGTCTATGAGAAACCATCTTGATCTCCAGAGACT	2845
241 CAATGGCTGTAAGCTTCTGTGCCAGGTCTATGAGAAACCATCTTGATCTCCAGAGACT	300
2850 GAGGTGGTAAAGTGTATCTCACCGCCGAGCTGTGAGAGTGTGCTGTGATCTC	2905
301 GAGGTGGTAAAGTGTATCTCACCGCCGAGCTGTGAGAGTGTGCTGTGATCTC	360
2910 GAGGAGCAGACCTGTAAGGCTGATCTCAGCGACAAATCCCTGTGGTACGTTGGGT	2965

Db 361 GAGGACAGACACCTGAAGAGCCCTGATCTACGAGCAATGCCCTGGTGAAGGAGGT 420  
QY 2970 TGCTGCGCTGTGCGAGGAGCTGAAGCAAAAGAGTGTCTGAGAGACTCGGGTTGA 3029  
Db 421 TGCTGCGCTGTGCGAGGAGCTGAAGCAAAAGAGTGTCTGAGAGACTCGGGTTGA 480  
QY 3030 GGCATGTGAGTGAATCTGTGATTTGCTGTGAGGACATCTCTGCGCTTCTGCAACCG 3089  
Db 481 GGCATGTGAGTGAATCTGTGATTTGCTGTGAGGACATCTCTGCGCTTCTGCAACCG 540  
QY 3090 GCATGTGACGAGTGAACCTGTGTGCAAGATTAATCTTCACTCCCAAGAGATGATGACT 3149  
Db 541 GCATGTGACGAGTGAACCTGTGTGCAAGATTAATCTTCACTCCCAAGAGATGATGACT 600  
QY 3150 GGTTCGCGCTTGTGCTGTGCGAGCTGAATTAATGAGGCGTGTGAATGCA 3209  
Db 601 GGTTCGCGCTTGTGCTGTGCGAGCTGAATTAATGAGGCGTGTGAATGCA 660  
QY 3210 GTACCTGTGCAATTAAGAGAGCTGTGAGAGAGTGAAGTCAAGCCCGAGTGT 3269  
Db 661 GTACCTGTGCAATTAAGAGAGCTGTGAGAGAGTGAAGTCAAGCCCGAGTGT 720  
QY 3270 AATTGACGCTGATGCTGCTTCTTTGATGAAGATGACCG 3308  
Db 721 AATTGACGCTGATGCTGCTTCTTTGATGAAGATGACCG 759  
RESULT 14  
US-10-027-632-258159  
; Sequence 258159, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR FILING DATE: 2002-07-12  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 258159  
; LENGTH: 2099  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-258159  
Query Match 4.3%; Score 174; DB 13; Length 2099;  
Best Local Similarity 100.0%; Pred. No. 7.4e-80;  
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 338 AGACCTGTGTGCCAAGCTGAGGATCCCACTGCAAGATACAGACCTGATGT 391  
RESULT 15  
US-10-027-632-258159  
; Sequence 258159, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 258159  
; LENGTH: 2099  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-258159  
Query Match 4.3%; Score 174; DB 17; Length 2099;  
Best Local Similarity 100.0%; Pred. No. 7.4e-80;  
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: July 19, 2005, 12:03:16  
Job time : 2599 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2005, 01:37:21 ; Search time 12239 Seconds

(without alignments) 12549.173 Million cell updates/sec

Title: US-10-066-521-5

Perfect score: 4035  
Sequence: 1 atggaagagagacaatcgct.....tcgaggggctggtctcttaa 4035

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : EST:

1: gb\_esc1:\*  
2: gb\_esc2:\*  
3: gb\_hnc:\*  
4: gb\_esc3:\*  
5: gb\_esc4:\*  
6: gb\_esc5:\*  
7: gb\_esc6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	679	16.8	791	7	CK000513 AGENCOURT
2	322	8.0	693	5	BU634350 UI-H-FIL-
3	299	7.4	684	5	BU630481 UI-H-FLO-
4	256	6.3	642	5	BU618831 UI-H-FIL-
5	116	2.9	526	8	AO542616 RPT-11-3
6	60	1.5	523	8	AO428788 CITR-EL-
7	58	1.4	814	6	CB228956 AGENCOURT
8	50	1.2	743	9	AG117567 Pan tlog1
9	24	0.6	917	2	BF133827 Reverse s
10	23	0.6	827	7	CR182764 Reverse s
11	22	0.5	308	7	CO111900 EK227159
12	22	0.5	392	7	BE406212 WHE0408-e
13	22	0.5	394	7	CF916020 B0986G02-
14	22	0.5	407	6	CB807101 AMGNNUC:S
15	22	0.5	425	7	BE406757 WHE0406-e
16	22	0.5	511	7	CF173704 B0927G04-
17	22	0.5	533	6	CA559680 K0261H12-
18	22	0.5	539	2	AW948258 RCO-MT001
19	22	0.5	546	6	CA561388 K0286C11-
20	22	0.5	546	7	CK546518 rwbnd0_01
21	22	0.5	559	6	CA561556 K0288H05-
22	22	0.5	583	6	CA559979 K0266B01-
23	22	0.5	585	8	AO515975 HS 5240 A
24	22	0.5	589	6	CA559886 K0264H01-

25	22	0.5	593	6	CA558845	CA558845 K0250F09-
26	22	0.5	599	7	CF915731	CF915731 B0984E03-
27	22	0.5	605	7	CK516658	CK516658 rwbnd0_00
28	22	0.5	625	6	CD338591	CD338591 ScFv537.
29	22	0.5	636	7	CF914959	CF914959 B0971H02-
30	22	0.5	643	7	CF913803	CF913803 B0954D05-
31	22	0.5	666	1	AV367637	AV367637 AV367637
32	22	0.5	719	7	CO806156	CO806156 AGENCOURT
33	22	0.5	768	7	CO808001	CO808001 AGENCOURT
34	22	0.5	798	2	BE643252	BE643252 Ct12_8_F0
35	22	0.5	835	4	BI957125	BI957125 HVSME000
36	22	0.5	1096	4	BC324013	BC324013 602422255
37	22	0.5	1119	4	BM454123	BM454123 AGENCOURT
38	22	0.5	1913	3	AK016782	AK016782 Mus muscu
39	22	0.5	3475	3	AK087774	AK087774 Mus muscu
40	22	0.5	395	1	AA316487	AA316487 EST188400
41	21	0.5	412	6	CB126685	CB126685 K-EST0175
42	21	0.5	419	7	CN265231	CN265231 170006000
43	21	0.5	444	2	AW439360	AW439360 xtl1h09.x
44	21	0.5	456	7	CO380162	CO380162 FRA3426.S
45	21	0.5	459	1	AI056454	AI056454 oy77a10.x

## ALIGNMENTS

RESULT 1  
LOCUS CK000513  
DEFINITION AGENCOURT 16368905 NIH MGC 221 Homo sapiens cDNA clone  
IMAGE:30708637 5', mRNA sequence.  
ACCESSION CK000513  
VERSION CK000513.1 GI:38526547  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLES  
JOURNAL  
COMMENT  
Unpublished (1999)  
Contract: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cga@bbs-remail.nih.gov  
Tissue procurement: James Martin, University of Iowa

cDNA Library Preparation: M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: NDAM1075 row: c column: 14  
High quality sequence stop: 688.  
Location/Qualifiers  
1. 791

## FEATURES

source  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30708637"  
/lab\_host="DH10B Tona"  
/clone\_id="NIH\_MGC\_221"  
/note="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;  
Site 2: NotI; Library is oligo-dT primed and directionally  
cloned. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to cDNA size fraction, ligated with  
EcoR I adaptor, digested with Not I and then cloned  
directionally into pYX-Asc vector. Average insert size  
4-5kb. Adaptors 5' (AATTGGCAGCAGG)3' and 5' d

(CCGCGGCGG)3', 3' linker sequence - GGGGCGCGCTGAGAGCC T18  
Sequencing primers 3' end: T3 promoter primer 5'd  
ATTACCCCTCAAGGGA)3', 5' End: T7 promoter primer 5'd  
(TATGACACTCAATAGG)3', library was constructed in the  
laboratory of M. Bento Soares. Note: this is a NIH\_MGC  
library"

```
RESULT 3
BU630481/c      684 bp  mRNA  linear  EST 23-SEP-2002
LOCUS           UI-H-F10-bdk-c-22-0-UI.s1 NCI CGAP F10 Homo sapiens cDNA clone
DEFINITION      UI-H-F10-bdk-c-22-0-UI 3', mRNA sequence.
ACCESSION       BU630481
VERSION         BU630481.1 GI:23297486
KEYWORDS        EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE       1 (bases 1 to 684)
AUTHORS         NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                Tumor Gene Index
                Unpublished (1997)
JOURNAL         Contact: Robert Strausberg, Ph.D.
COMMENT         Email: cgapbs-remail.nih.gov
                Tissue Procurement: James Martin
                cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
                DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                Clone Distribution: Clone distribution information can be obtained
                from Dr. M. Bento Soares, bento-soares@uiowa.edu
                Seq primer: M13 FORWARD
                POLYA=Yes.

FEATURES
    source
        1..684
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="UI-H-F10-bdk-c-22-0-UI"
            /tissue_type="Cell lines"
            /dev_stage="Adult"
            /lab_host="DH10B (Life Technologies)"
            /clone_lib="NCI CGAP F10"
            /note="Organ: Chondrosarcoma; Vector: pTT73-Pac
            (Pharmacia) with a modified polylinker; Site 1: EcoR I;
            Site 2: Not I; NCI CGAP F10 is a cDNA library derived from
            a pool of mRNA obtained from 4 cell lines from grade III
            chondrosarcoma tissues. The library was constructed
            according to Bonaldo, Lennon and Soares, Genome Research,
            6:791-806, 1996. First strand cDNA synthesis was primed
            with an oligo-dT primer containing a Not I site. Double
            stranded cDNA was ligated to an EcoR I adaptor, digested
            with Not I, and cloned directionally into pTT73-Pac
            vector. The oligonucleotide used to prime the synthesis of
            first-strand cDNA contains a library tag sequence that is
            located between the Not I site and the (dT)18 tail. The
            sequence tag for this library is GAGGTCGGTG. The cell line
            was provided by Dr James Martin from University of Iowa.
            TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix
            TAG_LIB=UI-H-F10
            TAG_SEQ=GAGTCGGTG"
```

```
Db 538 ACCGGCATCTGACACAGCTAAACCTGTGAGAAATTACTCAGTCCAAAGAAATGATGA 479
Qy 3146 AGCTGTGTTGGCCCTTTCCTGCTCCAGCTTAATTAATTTGGCTGTGGAAAT 3205
Db 478 AGCTGTGTTGGCCCTTTCCTGCTCCAGCTTAATTAATTTGGCTGTGGAAAT 419
Qy 3206 GGCAGTACCTGTCGAATTAAGAAAGCTGTGAGGAAGTCACTCAAGCCCCCA 3264
Db 418 GGCAGTACCTGTCGAATTAAGAAAGCTGTGAGGAAGTCACTCAAGCCCCCA 360

RESULT 4
BU618831/c      642 bp  mRNA  linear  EST 23-SEP-2002
LOCUS           UI-H-FH1-bf1-1-19-0-UI.s1 NCI CGAP FH1 Homo sapiens cDNA clone
DEFINITION      UI-H-FH1-bf1-1-19-0-UI 3', mRNA sequence.
ACCESSION       BU618831
VERSION         BU618831.1 GI:23285046
KEYWORDS        EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE       1 (bases 1 to 642)
AUTHORS         NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                Tumor Gene Index
                Unpublished (1997)
JOURNAL         Contact: Robert Strausberg, Ph.D.
COMMENT         Email: cgapbs-remail.nih.gov
                Tissue Procurement: James Martin
                cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
                DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                Clone Distribution: Clone distribution information can be obtained
                from Dr. M. Bento Soares, bento-soares@uiowa.edu
                Seq primer: M13 FORWARD
                POLYA=Yes.

FEATURES
    source
        1..642
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="UI-H-FH1-bf1-1-19-0-UI"
            /tissue_type="Cell line"
            /dev_stage="Adult"
            /lab_host="DH10B (Life Technologies)"
            /clone_lib="NCI CGAP FH1"
            /note="Organ: Chondrosarcoma; Vector: pTT73-Pac
            (Pharmacia) with a modified polylinker; Site 1: EcoR I;
            Site 2: Not I; NCI CGAP FH1 is a normalized cDNA library
            obtained from a cell line derived from grade I
            chondrosarcoma tissue. The library was constructed and
            normalized according to Bonaldo, Lennon and Soares, Genome
            Research, 6:791-806, 1996. First strand cDNA synthesis was
            primed with an oligo-dT primer containing a Not I site.
            Double stranded cDNA was ligated to an EcoR I adaptor,
            digested with Not I, and cloned directionally into
            pTT73-Pac vector. The oligonucleotide used to prime the
            synthesis of first-strand cDNA contains a library tag
            sequence that is located between the Not I site and the
            (dT)18 tail. The sequence tag for this library is
            AGAATCCGGC. The cell line was provided by Dr. James Martin
            from the University of Iowa.
            TAG TISSUE=Human Chondrosarcoma Cell Line C58 - Grade 1
            Chondrosarcoma
            TAG_LIB=UI-H-FH1
            TAG_SEQ=AGAATCCGGC"
```

```
ORIGIN
Query Match      6.3%; Score 256; DB 5; Length 642;
Best Local Similarity 100.0%; Pred. No. 1,7e-124; Indels 0; Gaps 0;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Oy	2516	AGAGCTGTGGGATACAGCCACGGGTTGCCAGAGTCTGGCCTCACGCCCTCGTCAGCAAC	25715
Dd	302	AGAGCTGTGGCATACAGCCACGGGTTGCCAGAGTCTGGCCTCACGCCCTCGTCAGCAAC	243
Oy	2576	GGAGCTTGACACACTGTGCCTTATCCAAACAACAGCTTGGGGAGCGAAGTGTAAAT	2631
Dd	242	GGAGCTTGACACACTGTGCCTTATCCAAACAAGCCTGGGGAGCGAAGTGTAAAT	187
RESULT 6			
AQ428788			
DEFINITION	CITRBI-EI-2560117.TF CITRBI-EI Homo sapiens genomic clone 2560117,	523 bp DNA linear GSS 24-MAR-1999	
ACCESSION	AQ428788		
VERSION	AQ428788.1	GI:4496554	
KEYWORDS	GSS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 523) Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,W. and Venter,J.C. Use of BAC End Sequences from Caltech Libraries for Sequence-Ready Map Building Unpublished (1997) Other-GSSs: CITRBI-EI-2560117.TR Contact: Shaying Zhao, William Nierman, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel.: 301 838 0200 Fax: 301 838 0208 Email: hbe@tigr.org Clones are available from Reseach Genetics (info@resgen.com). BAC end search page: <a href="http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html">http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html</a> . Seq primer: M13-21 Class: BAC ends.		
FEATURES			
source	Location/Qualifiers		
	1..523	/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		/clone="2560117"	
		/sex="male"	
		/cell_type="sperm"	
		/clone_id="CITRBI-EI"	
		/note="vector: pBELOBAC11; Site_1: EcoRI; Site_2: EcoRI; Caltech Human BAC Library D"	
ORIGIN			
Query Match	1.5%; Score 60; DB 8; Length 523;		
Best Local Similarity	100.0%; Pred. No. 1.3e-19;		
Matches	60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Oy	345	AGAAATTTCACAACTATGGAAACAAGAAGTGGCCACAGCAGACAGACAGAAACAAGG	404
Dd	21	AGAAATTTCACAACTATGGAAACAAGAAGTGGCCACAGCAGACAGAAACAAGG	80
RESULT 7			
LOCUS	CB228956		
DEFINITION	CB228956	814 bp mRNA linear EST 10-FEB-2003	
ACCESSION	CB228956		
VERSION	CB228956.1	GI:28280534	
KEYWORDS	EST.		
SOURCE	Macaca mulatta (rhesus monkey)		
ORGANISM	Macaca mulatta		

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cerothecidae; Macaca.

## REFERENCE

1 (bases 1 to 814)  
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

## AUTHORS

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

## JOURNAL

Unpublished (1997)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Dr. Elliot Spindel

CDNA Library Preparation: CLOTECH

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1CM3135 row: b column: 23

High quality sequence stop: 235.

## FEATURES

Location/Qualifiers

1..814

/organism="Macaca mulatta"

/mol\_type="mRNA"

/db\_xref="taxon:9544"

/clone="IMAGE:5884760"

/tissue\_type="Ovary"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NICHD\_Rh\_Ov1"

/note="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I; Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.0-4.0 kb. Tissue pooled from pre-pubertal, post pubertal sn menopausal monkeys. Constructed by Clontech. Note: this is a NICHD Library."

## ORIGIN

Query Match

Best Local Similarity 1.4%; Score 58; DB 6; Length 814;

Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## ORIGIN

Query

2289

GATGCGGTGTGAAGCCTTAAACACCCAAATGTTGTGAGCTTTGAGCGTGGAT 2346

140 GATGCGGTGTGAAGCCTTAAACACCCAAATGTTGTGAGCTTTGAGCGTGGAT 197

## Db

RESULT 8

AG117567/c

LOCUS

AG117567

DEFINITION

AG117567

ACCESSION

AG117567.1

VERSION

AG117567.1

KEYWORDS

GSS.

SOURCE

Pan troglodytes (chimpanzee)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.

REFERENCE

1

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

BAC end sequences of library PTB

Unpublished

2 (bases 1 to 743)

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

Direct Submission

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suhiro-chou, Tsukuba, Ibaraki, Japan

(E-mail: [chumpease@riken.go.jp](mailto:chumpease@riken.go.jp), URL: <http://hgp.gsc.riken.go.jp/>, Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS

COMMENT

JOURNAL

TITLE

AUTHORS

REFERENCE

1

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

BAC end sequences of library PTB

Unpublished

2 (bases 1 to 743)

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

Direct Submission

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suhiro-chou, Tsukuba, Ibaraki, Japan

(E-mail: [chumpease@riken.go.jp](mailto:chumpease@riken.go.jp), URL: <http://hgp.gsc.riken.go.jp/>, Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS

Sequencing: -21M13

LIBRARY

Vector : PKS145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1..743

/organism="Pan troglodytes"

/mol\_type="genomic DNA"

/db\_xref="taxon:9598"

/clone="PTB-125H04.F"

/sex="male"

/cell\_type="lymphoblast"

/clone\_lib="PTB Chimpanzee Male BAC library"

## ORIGIN

Query Match

Best Local Similarity 1.2%; Score 50; DB 9; Length 743;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## ORIGIN

Query

3332

GCTCTCTGAAAGCCGCGCATGCGCATGCTCTCTCTGCGGATGAC 3381

Db

296

GCTCTCTGAAAGCCGCGCATGCGCATGCTCTCTCTGCGGATGAC 247

## RESULT 9

BF133827/c

LOCUS

6017875291

DEFINITION

6017875291

ACCESSION

BF133827

VERSION

BF133827

KEYWORDS

EST.

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriongnathi; Muridae; Murinae; Mus.

REFERENCE

1

NIH-MGC <http://mhc.nci.nih.gov/>.

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM9238 row: h column: 13

High quality sequence stop: 630.

Location/Qualifiers

1..917

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="IMAGE:4006716"

/tissue\_type="tumor, metastatic to mammary"

/lab\_host="DH10B"

/clone\_lib="NCI-CCGAP Lu30"

/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; transgenic model MMT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

Investigator providing samples: Gilbert Smith, NIH

PRIMERS

COMMENT

JOURNAL

TITLE

AUTHORS

REFERENCE

1

NIH-MGC <http://mhc.nci.nih.gov/>.

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

```

Qy      1895 CCTTCTGCTCCAGCACTGTCCGT 1918
      |||||||
Db      742 CCTTCTGCTCCAGCACTGTCCGT 719

RESULT 10
LOCUS   CR182764      827 bp    DNA        linear    GSS 06-JUL-2004
DEFINITION Reverse strand read from insert in 5'HPT insertion targeting and
            chromosome engineering clone MHPN374002, genomic survey sequence.
ACCESSION CR182764
VERSION   CR182764.1 GI:49961613
KEYWORDS  GSS; genome survey sequence; MICR.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 827)
REFERENCE Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
AUTHORS Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
TITLE     Direct Submission
JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICR
FEATURES
            source
            1..827
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /clone="MHPN374002"
            /clone_lib="MHPN"

ORIGIN
Query Match      0.6%; Score 23; DB 9; Length 827;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2977 CTGTGCGAGGAGCTAGCAAAA 2999
      |||||||
Db      49 CTGTGCGAGGAGCTAGCAAAA 71

RESULT 11
LOCUS   CO311900/c      308 bp    mRNA        linear    EST 28-JUN-2004
DEFINITION EK227159.Sprime Exelixis FlyTag CK01 pCDNA-SK+ Drosophila
            melanogaster cDNA clone EK227159 5, mRNA sequence.
ACCESSION CO311900
VERSION   CO311900.1 GI:49366701
KEYWORDS  EST.
SOURCE    Drosophila melanogaster (fruit fly)
ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
            1 (bases 1 to 308)
REFERENCE Kopczynski,C., Platt,D., Campbell,J., Muzong,C., Lafter,A.,
AUTHORS Peterson,E. and Swimmer,C.
TITLE     Exelixis FlyTag EST Project CK01 Library
JOURNAL   Unpublished (2004)
COMMENT   Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: EK-2271 row: B column: 11
            High quality sequence stop: 267.
FEATURES
            source
            1..308
            /organism="Drosophila melanogaster"
            /mol_type="mRNA"
            /db_xref="taxon:7227"

```

```

ORIGIN
Query Match      0.5%; Score 22; DB 7; Length 308;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2553 GGCTCAGCCCTCGTCAGCAAC 2574
      |||||||
Db      201 GGCTCAGCCCTCGTCAGCAAC 180

RESULT 12
LOCUS   BE406212      392 bp    mRNA        linear    EST 21-JUL-2000
DEFINITION WHE0408_e11_e12B wheat etiolated seedling root cDNA library
            Triticum aestivum cDNA clone WHE0408_e11_e11, mRNA sequence.
ACCESSION BE406212
VERSION   BE406212.1 GI:9365680
KEYWORDS  EST.
SOURCE    Triticum aestivum (bread wheat)
ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooidae; Triticeae; Triticum.
            1 (bases 1 to 392)
REFERENCE Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D.,
AUTHORS Han,P.S., Heia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
            Seaton,C.L. and Tong,J.C.
TITLE     The structure and function of the expressed portion of the wheat
            genomes
JOURNAL   Unpublished (2000)
COMMENT   Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oanderson@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Strategene pBluescript SK reverse primer.
FEATURES
            source
            1..392
            /organism="Triticum aestivum"
            /mol_type="mRNA"
            /cultivar="Chinese Spring"
            /db_xref="taxon:4565"
            /clone="WHE0408_e11_e11"
            /tissue_type="Root"
            /dev_stage="Five day old etiolated seedling"
            /lab_host="B. coli 50LR"
            /clone_lib="Wheat etiolated seedling root cDNA library"
            /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
            Site_1: EcoRI; Site_2: XhoI; Seeds were
            surface-sterilized, germinated and grown aseptically in
            the dark at room temperature on filter paper with water,
            nystatin and cefotaxime in covered crystallization
            dishes. Roots were harvested. The tissue, total RNA, and
            poly(A) RNA were prepared, a cDNA library was made, and
            the cDNA clones were in vivo excised to give plusescript
            phagemids in the T7 Close lab (Choi, Close, Fenton) at the
            University of California, Riverside. Plasmid DNA
            preparations and DNA sequencing were performed in the OD
            Anderson lab (all other authors)."
ORIGIN
Query Match      0.5%; Score 22; DB 2; Length 392;

```



Best Local Similarity 100.0%; Pred. No. 28;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1482 TTCCAGAGACTTCTGTGCCGCC 1503

Db 162 TCTCCAGAGACTTCTGTGCCGCC 183

## RESULT 13

CF916020 394 bp mRNA linear EST 05-NOV-2003

LOCUS B0988G02-5 NIA Mouse Unfertilized Egg cDNA Library (Long 1) Mus

DEFINITION musculus cDNA clone NIA:B0988G02 IMAGE:30480745 5', mRNA sequence.

ACCESSION CF916020

VERSION CF916020.1 GI:38187222

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Piao, Y., Ko, N.-I., Lim, M. K. and Ko, M. S. H.

TITLE Construction of long-transcript enriched cDNA libraries from

submicrogram amounts of total RNAs by a universal PCR amplification

method

Genome Res. 11 (9), 1553-1558 (2001)

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@igsun.grc.nia.nih.gov

Plate: B0988 row: G column: 02

Seq primer: M13 Reverse

High quality sequence stop: 394

## FEATURES

## source

Location/Qualifiers  
1..394  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/db\_xref="taxon:B0988G02-5"  
/clone="NIA:B0988G02 IMAGE:30480745"  
/dev\_stage="Unfertilized Egg"  
/lab\_host="DH10B"  
/clone\_1ib="NIA Mouse Unfertilized Egg cDNA Library (Long 1)"  
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). (PMID: 11544199)). Total RNAs were extracted from a pool of 1468 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen):  
5'-PACATAGTTCTAGATCGAGCGCGCCCTTTTCTTTT-3',  
treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker Lb-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Tag polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao."

## ORIGIN

Query Match 0.5%; Score 22; DB 7; Length 394;

Best Local Similarity 100.0%; Pred. No. 28;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1417 TTTCATGATGACATCTCTCC 1438

Db 56 TTTCATGATGACATCTCTCC 77

## RESULT 14

CB807101 407 bp mRNA linear EST 16-MAY-2003

LOCUS AMGNNUC:SRPB2-00129-B10-A srpb2 (10220) Rattus norvegicus cDNA

DEFINITION clone srpb2-00129-b10 5', mRNA sequence.

ACCESSION CB807101

VERSION CB807101.1 GI:29925806

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Rattus.

TITLE 1 (bases 1 to 407)

JOURNAL Amgen EST Program

COMMENT Unpublished (2003)

Contact: Dan Fitzpatrick

Amgen, Inc

One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00129 row: b column: 10.

Location/Qualifiers

1..407

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

/clone="srpb2-00129-b10"

/tissue\_type="prostate tissue"

/clone\_1ib="srpb2 (10220)"

/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; rat prostate normalized double selected poly(A+) mRNA size fraction > 1 kb"

## ORIGIN

Query Match 0.5%; Score 22; DB 6; Length 407;

Best Local Similarity 100.0%; Pred. No. 28;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 482 TTGAACACAGCTGCTGACTG 503

Db 167 TTGAACACAGCTGCTGACTG 188

## RESULT 15

BE406757 425 bp mRNA linear EST 21-JUL-2000

LOCUS WHB0406 e11 e11s wheat etiolated seedling root cDNA library

DEFINITION Triticum aestivum cDNA clone WHB0406\_e11\_e11, mRNA sequence.

ACCESSION BE406757

VERSION BE406757.1 GI:9366225

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D.,

Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rauech, C.J.,

Seaton, C.L. and Tong, J.C.

TITLE The structure and function of the expressed portion of the wheat

genomes

JOURNAL  
COMMENT

Unpublished (2000)  
 Contact: Olin Anderson  
 US Department of Agriculture, Agriculture Research Service, Pacific  
 West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105595773  
 Fax: 5105595818  
 Email: oanderson@pw.usda.gov  
 Sequence have been trimmed to remove vector sequence and low  
 quality sequence with phred score less than 20  
 Seq primer: Strategene SK primer.

## FEATURES

source

1. .425  
 Location/Qualifiers  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Chinese Spring"  
 /db\_xref="taxon:4565"  
 /clone="WHE0406\_e11\_e11"  
 /tissue\_type="Root"  
 /dev\_stage="Five day old etiolated seedling"  
 /lab\_host="E. coli SOLR"  
 /clone\_lib="Wheat etiolated seedling root cDNA library"  
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;  
 Site 1: EcoRI; Site 2: XhoI; Seeds were  
 surface-sterilized, germinated and grown aseptically in  
 the dark at room temperature on filter paper with water,  
 nystatin and cefotaxime in covered crystallization  
 dishes. Roots were harvested. The tissue, total RNA, and  
 poly(A) RNA were prepared, a cDNA library was made, and  
 the cDNA clones were in vivo excised to give Bluescript  
 phagemids in the TJ Clonase lab (Choi, Close, Fenton) at the  
 University of California, Riverside. Plasmid DNA  
 preparations and DNA sequencing were performed in the OD  
 Anderson lab (all other authors)."

## ORIGIN

Query Match 0.5%; Score 22; DB 2; Length 425;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1482 TCTCCAGGACTTCTGTGCGGCC 1503  
 ||||||||||||||||||||  
 Db 117 TCTCCAGGACTTCTGTGCGGCC 138

Search completed: July 19, 2005, 11:09:20  
 Job time : 12248 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2005, 14:43:30 ; Search time 17004 Seconds  
(without alignments)  
11498.283 Million cell updates/sec

Title: US-10-066-521-5  
Perfect score: 4035  
Sequence: 1 atggaagagagaacaatcgct.....tcgaggcgctggtcctaa 4035

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 2422767955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_srs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3172.6	78.6	3885	9	AY154460 Homo sapi
2	3166.2	78.5	3885	9	AY054986 Homo sapi
3	3166.2	78.5	3900	6	AX427610 Sequence
4	3139	77.8	3830	6	AX704823 Sequence
5	3133.6	77.7	3489	6	AX478549 Sequence
6	3133.6	77.7	3926	6	AX704821 Sequence
7	3057.8	75.8	5859	6	AX459881 Sequence
8	3057.8	75.8	6939	6	AX459873 Sequence
9	3057.8	75.8	6939	6	AX459891 Sequence
10	2553.8	63.3	2753	6	C0731113 Sequence
11	1598.8	39.6	157141	9	AC011470 Homo sapi
12	1598.8	39.6	193609	9	AC024580 Homo sapi
13	1566	38.8	3545	4	AY721594 Bos tauru
14	1287.8	31.9	3405	10	AY329487 Mus muscu
15	1286.2	31.9	3432	10	AY329484 Mus muscu
16	1282.8	31.8	3480	10	AY196361 Mus muscu
17	1281.2	31.8	3470	10	AY196362 Mus muscu
18	1279.6	31.7	3453	10	AY329486 Mus muscu
19	1278	31.7	3447	6	AX427592 Sequence

20	1278	31.7	3447	10	AF074018 Mus muscu
21	1278	31.7	3454	10	AY329490 Mus muscu
22	1276.4	31.6	3534	10	BC053384 Mus muscu
23	1152.8	28.6	1157	6	AX427588 Sequence
24	1144	28.4	3375	10	AY329489 Mus muscu
25	1140.8	28.3	3372	10	AY329485 Mus muscu
26	1008.6	25.0	3303	10	AY329488 Mus muscu
27	1005.4	24.9	3303	10	AY329491 Mus muscu
28	988.6	24.5	167509	2	AC012107 Homo sapi
29	840.4	20.8	193662	2	AC127352 Mus muscu
30	840.4	20.8	213424	2	AC148980 Mus muscu
31	840.4	20.8	299377	2	AC127238 Mus muscu
32	756.6	18.8	1075	6	AX427590 Sequence
33	729.4	18.1	3282	9	AY154469 Homo sapi
34	649.4	16.1	231371	2	AC132633 Rattus no
35	649.4	16.1	323826	2	AC110429 Rattus no
36	645.2	16.0	1576	10	AF143565 Mus muscu
37	528.4	13.1	167509	2	AC012107 Homo sapi
38	430.4	10.7	2964	10	AY596199 Mus muscu
39	428.8	10.6	3281	10	AY673647 Mus muscu
40	422.4	10.5	3190	6	AX714395 Sequence
41	422.4	10.5	3190	9	AK056688 Homo sapi
42	422.4	10.5	3333	9	AF442488 Homo sapi
43	422.4	10.5	3339	9	AY072792 Homo sapi
44	422.4	10.5	3368	9	AF479747 Homo sapi
45	419.2	10.4	2575	6	AX320139 Sequence

## ALIGNMENTS

RESULT 1  
AY154460  
LOCUS AY154460 3885 bp mRNA linear PRI 20-FEB-2003  
DEFINITION Homo sapiens NALP5 (NALP5) mRNA, complete cds.  
ACCESSION AY154460  
VERSION AY154460.1 GI:28436363  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 (bases 1 to 3885)  
Techopp,J., Martinon,F. and Burns,K.  
NALP5: a novel protein family involved in inflammation  
JOURNAL Nat. Rev. Mol. Cell Biol. 4 (2), 95-104 (2003)  
MEDLINE 22451042  
PUBMED 12563287

REFERENCE  
2 (bases 1 to 3885)  
Martinon,F., Hotmann,K. and Techopp,J.  
Direct Submission  
Submitted (25-SEP-2002) Institute of Biochemistry, University of Lausanne, ch. des Boveresses 155, Epalinges, VD 1066, Switzerland

## FEATURES

source location/Qualifiers  
1..3885  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..3885  
/gene="NALP5"  
1..3603  
/gene="NALP5"  
/note="member of the NALP protein family involved in inflammation; contains NACHT, LRR and PYD domains; similar to mouse MATER protein"  
/codon\_start=1  
/product="NALP5"  
/protein\_id="AA018156.1"  
/db\_xref="GI:28436364"  
/translation="MKVAGGIELGAAALSLSPALVTLSTGPTCSILPKNPDPONTSSOPCKMEGDKSLFSSYGLWGCIVLEDKKEPOTFEKLKKSSSTGTCISIPQFEINAVEECALILHRYVAGSLAIAVTSISIFENNMLRTLSKPADDKRRHSPEDPATMTDGGPSKEKVPGISQAVQDSATRAETKEQELISQAMEBSGATLAETBEQELISQAMEBSGA

TAATEBOGHGDTWYKSHVMTKFAEEDVRSFENTADMPMEMQTLAAGFSDNRWG  
FRPRTVYLHSGKSGKLSALARIYVLQMGQGLYOCMFYFLPREXOKKESSTVE  
FISREMPDQSA.PVTEIMRSPRELIPIIDGPDIGSVLNDPRKLCMAEXOPPTLIR  
SLRKVILPESFLIYTVRDVGTREKSKSEVSPRYLIVGISEGRIHLLRIGGEHO  
KIQGKALIMNREBELDQCVPAVSLICVALQLODVESVAPEFQITLIGLAAVPH  
QUTPRGVRCINLEERVLKRFCEMAVEGWNKRSVADGDLWQIGBSEBLRFLH  
NMLILPDSHCBEYTFPHSLQDPCALUYVLBGLIEIPALCPVEKTKSMELKQ  
GFHLSMKKEPFLVSEVDYRPLEVLIGCEVPLGVKOLHMLHWLISQOONATLP  
DILDAFHCLPETOKEFVRLANSPFOEVLPINONLDIASPFOHCGRYLRKIRVDY  
KGIIPRDSARACPVPLMRBDKTLIREQWDPSCMGTHTHLOLDIGSSILIBRAM  
KTLCAKLRHPKIOITLMPRNAQITPVGQHMRIYMARNTLSMLGTHILKEEDVRM  
ACEALRKHRCLESRLDCCGLTHACVYKISQILITTSLSLSIAGKVDYGLMPL  
SDALRVSQCALILIEDQGITATGCGSLAALVSNRSLTHLSINPENNQVLLCE  
RSMRLPSCLORLMNOCHLDTPAGCGFALALMNSMTLTHLSINPENNQVLLCE  
VMRSPCHLODELIVKCHLTAACCELSCVLSRSHKLSLITDNLALGDGVALLCEG  
LKOKNSVTRIGAKCGLTSPCCBLSLALSINRHTLSLNVONNPSKGMKCSAF  
ACPTSNLOIIGLMKWQIPVQIRKLEBVLKPRVLDGSHWSFDEDRYMKW"

## ORIGIN

Query Match 78.6%; Score 3172.6; DB 9; Length 3885;  
Best Local Similarity 96.3%; Pred. No. 0;  
Matches 3308; Conservative 0; Mismatches 4; Indels 123; Gaps 1;

QY 1 ATGAGAGAGACAAATCGCTCACTTTTCCAGCTACGGGCTGCAATGGTGTCTATGAG 60  
DB 154 ATGAGAGAGACAAATCGCTCACTTTTCCAGCTACGGGCTGCAATGGTGTCTATGAG 213  
QY 61 CTAGCAAGAGAAATTTTCAAGCATTTCAAGAAATTAATAAGAAATTTTCAAGATCG 120  
DB 214 CTAGCAAGAGAAATTTTCAAGCATTTCAAGAAATTAATAAGAAATTTTCAAGATCG 273  
QY 121 ACCACATGCTCTATTCACAGTTTGAATTCAGAAATGCAAGTGAATGTCTGGACATC 180  
DB 274 ACCACATGCTCTATTCACAGTTTGAATTCAGAAATGCAAGTGAATGTCTGGACATC 333  
QY 181 CTCTTGATGATATTAATGAGCATCGCTGGCTGGCTGATCGTCCATTAGCATCTTTGAA 240  
DB 334 CTCTTGATGATATTAATGAGCATCGCTGGCTGGCTGATCGTCCATTAGCATCTTTGAA 393  
QY 241 AACATGAACCTGCCGAACCTCTCGGAGAAAGGCAAGGATGACATGATAAA----- 289  
DB 394 AACATGAACCTGCCGAACCTCTCGGAGAAAGGCAAGGATGACATGATAAAACATTCACCA 453  
QY 290 ----- 289  
DB 454 GAAATCTCTGAAGCAACATGACTGACCAAGGACCAAGCAAGAAAAAGTCCAGGAATT 513  
QY 290 -----AATTTTCA 297  
DB 514 TCACAAAGCTGTGCAACAAGATAGTGCAAGCTGCAAGAGACAAAGAAATTTTCA 573  
QY 298 CAAGTATGAGAACAAAGAGTGCCACAGACGACAGACAAAGAAACAAAGAAATTTTCA 357  
DB 574 CAAGTATGAGAACAAAGAGTGCCACAGACGACAGACAAAGAAACAAAGAAATTTTCA 633  
QY 358 GCTATGAGAACAAAGAGTGCCACAGACGACAGACAAAGAAACAAAGCATGGAAGTAC 417  
DB 634 GCTATGAGAACAAAGAGTGCCACAGACGACAGACAAAGAAACAAAGCATGGAAGTAC 693  
QY 418 ACATGAGGACTACAGAGTCAAGTATGACCAATTCGCTGAGAGAGAGATGTAAGTCTGT 477  
DB 694 ACATGAGGACTACAGAGTCAAGTATGACCAATTCGCTGAGAGAGAGATGTAAGTCTGT 753  
QY 478 AGTTTGAACCACTGCTGCTGACTGCGCGGAATATGCAAGTGGCTGTGCTTTGAT 537  
DB 754 AGTTTGAACCACTGCTGCTGACTGCGCGGAATATGCAAGTGGCTGTGCTTTGAT 813  
QY 538 TCAGACCGGTGGGGCTTCCGGCTCGACAGGTGTTCTGACGGAATTCAGGAATTTGGG 597  
DB 814 TCAGACCGGTGGGGCTTCCGGCTCGACAGGTGTTCTGACGGAATTCAGGAATTTGGG 873  
QY 598 AAATCGGCTTAGCCAGAGAGATCGTGTGCTGGGCGCAAGGTGACTTACACAGGGA 657

DB 874 AAATCGGCTTAGCCAGAGAGATCGTGTGCTGGGCGCAAGGTGACTTACAGGGA 933  
QY 658 ATGTTCTCTACGCTTCTCTCTCCCGTTAGAGAGATGACGGGAAGAGAGAGAGAGT 717  
DB 934 ATGTTCTCTACGCTTCTCTCTCCCGTTAGAGAGATGACGGGAAGAGAGAGAGT 993  
QY 718 GTCAAGAGTTCACTCCAGAGAGTGGCCAGACTCCAGAGCTCCGGTACCGAGATCATG 777  
DB 994 GTCAAGAGTTCACTCCAGAGAGTGGCCAGACTCCAGAGCTCCGGTACCGAGATCATG 1053  
QY 778 TCCGACCCAGAAAGGCTGTTTTCATCATTTGACCGTTTCATGACCTGGGCTCTGTCTC 837  
DB 1054 TCCGACCCAGAAAGGCTGTTTTCATCATTTGACCGTTTCATGACCTGGGCTCTGTCTC 1113  
QY 838 AACATGACCAAAAGCTCTGCAAGACCTGGGCTGAGAGAGCGCTCGTTCAACCTCATA 897  
DB 1114 AACATGACCAAAAGCTCTGCAAGACCTGGGCTGAGAGAGCGCTCGTTCAACCTCATA 1173  
QY 898 CGCAGTCTGCTGAGAGAGTCTCTGCTCCCTGAGTCTTCTGATGTCACCGTACAGAGAC 957  
DB 1174 CGCAGTCTGCTGAGAGAGTCTCTGCTCCCTGAGTCTTCTGATGTCACCGTACAGAGAC 1233  
QY 958 GTGAGCACAGAAAGCTCAAGTACAGAGTGTGTCTCCCGTTACTGTTAGTAGAGA 1017  
DB 1234 GTGAGCACAGAAAGCTCAAGTACAGAGTGTGTCTCCCGTTACTGTTAGTAGAGA 1293  
QY 1018 ATCTCCGGGGAACAAAGATTCACCTTGCTCTTGAAGCGGGATTTGGTAGCATCAGAA 1077  
DB 1294 ATCTCCGGGGAACAAAGATTCACCTTGCTCTTGAAGCGGGATTTGGTAGCATCAGAA 1353  
QY 1078 ACACAAAGGTTGCTGCGATCATGACAAACCGTAGCTGCTGACACAGTGCAGAGTCC 1137  
DB 1354 ACACAAAGGTTGCTGCGATCATGACAAACCGTAGCTGCTGACACAGTGCAGAGTCC 1413  
QY 1138 GCCGTGGCTCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197  
DB 1414 GCCGTGGCTCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1473  
QY 1198 GCCCTTTCAACCAAGGCTCAGAGGCTGACGCGCTTTTGTGTTCAATCACTAC 1257  
DB 1474 GCCCTTTCAACCAAGGCTCAGAGGCTGACGCGCTTTTGTGTTCAATCACTAC 1533  
QY 1258 CCTGAGGCGTGTGCTGCGGCGCTGCTCAATCTGAGAGAAAGTGTCTGAAACGCTTC 1317  
DB 1534 CCTGAGGCGTGTGCTGCGGCGCTGCTCAATCTGAGAGAAAGTGTCTGAAACGCTTC 1593  
QY 1318 TGCCGTATGCTGAGAGAGTGTGAATGAGAAATCAAGTGTGATGAGTACGACCTC 1377  
DB 1594 TGCCGTATGCTGAGAGAGTGTGAATGAGAAATCAAGTGTGATGAGTACGACCTC 1653  
QY 1378 ATGTTTCAAGGACTCCGGGAGTCTGAGCTCCGCTCTGTTTCAATGAAACATCTCTC 1437  
DB 1654 ATGTTTCAAGGACTCCGGGAGTCTGAGCTCCGCTCTGTTTCAATGAAACATCTCTC 1713  
QY 1438 CCAGACAGCCACTGTGAGAGTACTACACCTTCTTCCACCTCACTGCTCCAGACCTTCTGT 1497  
DB 1714 CCAGACAGCCACTGTGAGAGTACTACACCTTCTTCCACCTCACTGCTCCAGACCTTCTGT 1773  
QY 1498 GCCGCTTTGTAATAAGTGTGAGAGGCTGGAATTCAGGCAAGCTCTGCGCTCTGTAC 1557  
DB 1774 GCCGCTTTGTAATAAGTGTGAGAGGCTGGAATTCAGGCAAGCTCTGCGCTCTGTAC 1833  
QY 1558 GTTGAAGAACAAAGAGTCTATGAGACTTAAACAGGCAAGCTTCAATATCACTCGCTT 1617  
DB 1834 GTTGAAGAACAAAGAGTCTATGAGACTTAAACAGGCAAGCTTCAATATCACTCGCTT 1893  
QY 1618 TGAATGAAGGTTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1677  
DB 1894 TGAATGAAGGTTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1953  
QY 1678 CTGCTGGGCTGTCCGTTTCCCTGAGGAGTGAAGCAGAGCTTCTGCACTGGGCTCTCTG 1737  
DB 1954 CTGCTGGGCTGTCCGTTTCCCTGAGGAGTGAAGCAGAGCTTCTGCACTGGGCTCTCTG 2013

1738 TTGGGTGAGACGCTAATGCAACCAACCCAGAGAGACACCTGAGGCTTCCACTGTCTT 1797  
 2014 TTGGGTGAGACGCTAATGCAACCAACCCAGAGAGACACCTGAGGCTTCCACTGTCTT 2073  
 1798 TTGAGACTCAAGACAAAGAGTTGTGTGCTTGGCAATTAACAGCTTCCAGAAAGTGTGG 1857  
 2074 TTGAGACTCAAGACAAAGAGTTGTGTGCTTGGCAATTAACAGCTTCCAGAAAGTGTGG 2133  
 1858 CTTCGATTAACCAAGACCTGATCTTGATAGATCTTCTTGTGCTTCCAGAGACTGTCTGG 1917  
 2134 CTTCGATTAACCAAGACCTGATCTTGATAGATCTTCTTGTGCTTCCAGAGACTGTCTGG 2193  
 1918 TATTGGCGGAAAATTCCGGGTGATGTCAAAGGATCTTCCCAAGAGATGAGCCGCTGAG 1977  
 2194 TATTGGCGGAAAATTCCGGGTGATGTCAAAGGATCTTCCCAAGAGATGAGCCGCTGAG 2253  
 1978 GCATGTCCTGTGTCCTCTATGATGCGGGATTAAGACCTTCATTTAGAGAGCATGGGAA 2037  
 2254 GCATGTCCTGTGTCCTCTATGATGCGGGATTAAGACCTTCATTTAGAGAGCATGGGAA 2313  
 2038 GATTTCTGCTCATGCTTGGACACCAACCAACCTGCGGAGCTGAGCCTGGGACAGCAGC 2097  
 2314 GATTTCTGCTCATGCTTGGACACCAACCAACCTGCGGAGCTGAGCCTGGGACAGCAGC 2373  
 2098 ATCCGAGAGAGCGGGCCATGAAGACCTGTGTGCGCAAGCTGAGGATCCCACTGCAAG 2157  
 2374 ATCCGAGAGAGCGGGCCATGAAGACCTGTGTGCGCAAGCTGAGGATCCCACTGCAAG 2433  
 2158 ATACAGACCTGATTTTGAAGATGACAGATTAACCCCTGTGTGTCAGACCTCTGAGAG 2217  
 2434 ATACAGACCTGATTTTGAAGATGACAGATTAACCCCTGTGTGTCAGACCTCTGAGAG 2493  
 2218 ATGCTGATGGCCCAACCGTAACCTTAAGATCCTTCAACTTGGAGAGCAACCACTGAAGAA 2277  
 2494 ATGCTGATGGCCCAACCGTAACCTTAAGATCCTTCAACTTGGAGAGCAACCACTGAAGAA 2553  
 2278 GAGAGATGAAGATGCGCGTGAAGCCCTTAACCAACCCAAATGTTTGTGAGAGCTTTG 2337  
 2554 GAGAGATGAAGATGCGCGTGAAGCCCTTAACCAACCCAAATGTTTGTGAGAGCTTTG 2613  
 2338 AGGCTGATGATGCTGTGATTTGACCCATGCTGTACTGAAAGATCTCCCAATCTTTAG 2397  
 2614 AGGCTGATGATGCTGTGATTTGACCCATGCTGTACTGAAAGATCTCCCAATCTTTAG 2673  
 2398 ACCTCCCGACGCTGAAATCTCTGAGCCTGCGAGAGAAACAAGGTGACAGACCAAGGAGTA 2457  
 2674 ACCTCCCGACGCTGAAATCTCTGAGCCTGCGAGAGAAACAAGGTGACAGACCAAGGAGTA 2733  
 2458 ATGCTCTCAGTGAATGCTTGAAGATCTTCCAGTGGCGCTTGCAGAGCTGATCTGAGAG 2517  
 2734 ATGCTCTCAGTGAATGCTTGAAGATCTTCCAGTGGCGCTTGCAGAGCTGATCTGAGAG 2793  
 2518 GACTGTGGCATCAAGCAAGCGGTTGCCAGAGTCTGGCCTGAGCCCTGTCAGCAACCGG 2577  
 2794 GACTGTGGCATCAAGCAAGCGGTTGCCAGAGTCTGGCCTGAGCCCTGTCAGCAACCGG 2853  
 2578 AGCTTGACACACTGTGCTTATCAACAACAAGCTGGGAGCAAGAGTGTAAATCTACTG 2637  
 2854 AGCTTGACACACTGTGCTTATCAACAACAAGCTGGGAGCAAGAGTGTAAATCTACTG 2913  
 2638 TGTGATCAATGAGGCTTCCCACTGTAGTGTGCAAGAGCTGATCTGTAATCAGTCCAC 2697  
 2914 TGTGATCAATGAGGCTTCCCACTGTAGTGTGCAAGAGCTGATCTGTAATCAGTCCAC 2973  
 2698 CTGAGCAGGCGTGGCTGTGTTTCTTGACATTTGGCTTATGAGGTAATCATAGGCTGAG 2757  
 2974 CTGAGCAGGCGTGGCTGTGTTTCTTGACATTTGGCTTATGAGGTAATCATAGGCTGAG 3033  
 2758 CACCTGAGCCTTATGATGAACCTGTGAGAGCAATGAGGCTGTAATCTGTGCGAGGTC 2817  
 3034 CACCTGAGCCTTATGATGAACCTGTGAGAGCAATGAGGCTGTAATCTGTGCGAGGTC 3093

2818 ATGAGAGAACCATCTTGTATCTCCAGAGACCTGAGATTGTGAAGTGTATCTACACGCC 2877  
 3094 ATGAGAGAACCATCTTGTATCTCCAGAGACCTGAGATTGTGAAGTGTATCTACACGCC 3153  
 2878 GCGTCTGTGAGAGTCTGTCTGTGTGATCTTGAGAGACAGACACTGAAGAGCTTGAT 2937  
 3154 GCGTCTGTGAGAGTCTGTCTGTGTGATCTTGAGAGACAGACACTGAAGAGCTTGAT 3213  
 2938 CTCACGGAACATGCGCTGTGAGAGCGTGGGTTGCTGTGCTGTGAGAGGACTGAAGCA 2997  
 3214 CTCACGGAACATGCGCTGTGAGAGCGTGGGTTGCTGTGCTGTGAGAGGACTGAAGCA 3273  
 2998 AAGAAGTGTGTGAGACGAGACTGCGGTTGAAGGATGAGACTGACTTGTGATTTGCTGT 3057  
 3274 AAGAAGTGTGTGAGACGAGACTGCGGTTGAAGGATGAGACTGACTTGTGATTTGCTGT 3333  
 3058 GAGGACTCTCTTGGCCCTTCTGCAACCGGACATCTGACAGCTTAACCTGTGAG 3117  
 3334 GAGGACTCTCTTGGCCCTTCTGCAACCGGACATCTGACAGCTTAACCTGTGAG 3393  
 3118 AATACTTCACTCCCAAGAGATGATGAAGCTGTGTTGGGCTTGGCTTCCACAGCT 3177  
 3394 AATACTTCACTCCCAAGAGATGATGAAGCTGTGTTGGGCTTGGCTTCCACAGCT 3453  
 3178 AACTTACAGATTAATGGCTGTGAGAAATGGAGTACCTCTGCAAAATGAAGAGCTGCT 3237  
 3454 AACTTACAGATTAATGGCTGTGAGAAATGGAGTACCTCTGCAAAATGAAGAGCTGCT 3513  
 3238 GAGAGAGTGAAGTACTCAAGCCCGAGTGTGAATTAAGAGTGTGACTTTTGTAT 3297  
 3514 GAGAGAGTGAAGTACTCAAGCCCGAGTGTGAATTAAGAGTGTGACTTTTGTAT 3573  
 3298 GAGAGTGAAGTACTCAAGCCCGAGTGTGAATTAAGAGTGTGACTTTTGTAT 3357  
 3574 GAGAGTGAAGTACTCAAGCCCGAGTGTGAATTAAGAGTGTGACTTTTGTAT 3631

RESULT 2  
 AY054986 3885 bp mRNA linear PRI 02-APR-2002  
 LOCUS  
 DEFINITION Homo sapiens maternal-antigen-that-embryos-require protein (MATER)  
 AY054986  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 (bases 1 to 3885)  
 Tong, Z.-B., Bondy, C.A., Zhou, J., and Nelson, L.M.  
 A human homologue of mouse Mater, a maternal effect gene essential  
 for early embryonic development  
 Hum. Reprod. 17 (4), 903-911 (2002)  
 JOURNAL  
 MEDLINE  
 PUBMED  
 11925379  
 2 (bases 1 to 3885)  
 REFERENCE  
 Tong, Z.-B., Bondy, C.A., and Nelson, L.M.  
 Direct Submission  
 JOURNAL  
 Submitted (30-AUG-2001) Developmental Endocrinology Branch, NICHD,  
 NIH, 10 Center Drive, Bethesda, MD 20892/1862, USA  
 FEATURES  
 source  
 1. 3885  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 1. 3885  
 /gene="MATER"  
 1. 3603  
 /gene="MATER"  
 /note="oocyte-specific protein; similar to Mus musculus  
 Mater encoded by GenBank Accession Numbers AF074018 and

gene

CDS

AF143559-AF143573"  
/codon\_start=1  
/product="maternal-antigen-that-embryos-require protein"  
/protein\_id="AA15549.1"  
/db\_xref="GI:19882273"  
/translation="MKVAGGLEAALASPRALVTLSTGPTCSILPKNLPQNLS  
SQPCIKMEGDKSLTPFSYGLQMCYIELDKEEPQTKELKKKSESTCSIPQETIE  
NANVELLALHEYYGASIAWATSISITENNRLTLISKADDMKRSDEPRLATMD  
QSPKEKIPGISOAVODSATATAEETKEISQAMEOGATAEETBOEISQAMEOGA  
TAETREOGHGSDPTWDYKSHYTKFAEEDRBSFENALADMPMOTIAGAPDSDRG  
FRPRTVLHGKSGIGKSALARKIVLCAHOGSLYQMFSTYFLFVREKORKESSVE  
FISREWDSQAPVTEIMSRPERLLFIIDGFDDLSVINDKFLCKMAEKQPPRLIR  
SLRKVLPESEFLIVTVADVGTKEKSVSPRYLVSGLEGRILHLLERGGHEHQ  
KTQGRRAIINRELDQCVPAVGSILVALQLODVESVAPNOCITLGHAAFLFH  
OLTEGVRRCLNLEERVYLKRFQMAVEGWNKSVKVDGDLVAVOGISEELFALFH  
MNLIPDSHCERYTTFRLSLQDFCALVYLEGLIEPALCPVYEKTRSMELKOA  
GFTHLSLMWRKFLGLVSEDEVRLPLEVLLGCPVLGVQKLLHWVSLGQDPNATPFG  
DTLDAFHLFETQKEFVRLALNSFOEWMPLINQNLDAISFCLQHPYRKLRVY  
KGIPEPDESAPCAVPLMWRDKTLIEBOMEDFCSMLDTHPHLQOLDGSSILTERAM  
KTLCAKHPPTCKIOTLMFRNAOITPGVOIMRYMAMRNLSINLGGTHLKEDVDM  
ACEALKHPKCLLESRLDDCGLTACVYKISQILITTSPLKSLAGNKVTDGVTPL  
SDALRVGOCALOKILIEDCGITATGCCSLAALVNSLTHLCLSNNSLGNENILIC  
RSMRLPHOSLQRLMNOCHDITAGCGSLALMGENSVLTHLSLSPNVEDGVKLLCE  
VMSRPSCHLQDLELVKCHLITVACCELSCVISRSRLKSLDTLNNALFDGVAAFCBG  
LKORNSVLTRLGACGLTSDCEBSLSTALSCNRHLISLNVONNFSPKMMKCSAF  
ACPTSNLQIIGLMKQYPVQIRKILLEVQLKPRVIVIDGWSHSEDEDRYWMKN"

## ORIGIN

Query Match 78.5%; Score 3166.2; DB 9; Length 3885;  
Best Local Similarity 96.2%; Pred. No. 0;  
Matches 3304; Conservative 0; Mismatches 8; Indels 123; Gaps 1;

1 ATGGAAGAGACAATTCGCTCACCCTTTCCAGCTACGGGGTGCATGTGTCATATAG 60  
Db 154 ATGGAAGAGACAATTCGCTCACCCTTTCCAGCTACGGGGTGCATGTGTCATATAG 213  
Qy 61 CTAGAACAAGAAATTTCAACATTTCAAGAAATTTCTAAAGAAATCTTTCAGAAATCG 120  
Db 214 CTAGAACAAGAAATTTCAACATTTCAAGAAATTTCTAAAGAAATCTTTCAGAAATCG 273  
Qy 121 ACCCATCTCTATTTCCACAGTTTGAATTCAGAAATGCCAACGTGGAATGTCTGGCACTC 180  
Db 274 ACCCATCTCTATTTCCACAGTTTGAATTCAGAAATGCCAACGTGGAATGTCTGGCACTC 333  
Qy 181 CTCTTGATGAGTATTAGAGACATCGTGGCTGGGCTAGTCCATAGACATCTTTGAA 240  
Db 334 CTCTTGATGAGTATTAGAGACATCGTGGCTGGGCTAGTCCATAGACATCTTTGAA 393  
Qy 241 AACATGAACCTGCGAACCTCTCGGAGAAAGCACGGATGACATGAAAA----- 289  
Db 394 AACATGAACCTGCGAACCTCTCGGAGAAAGCACGGATGACATGAAAAAGCATTCACCA 453  
Qy 290 ----- 289  
Db 454 GAAAGATCTTGAAGCAACGATGCTGACCAAGGACCAAGCAAGAAAAAGTGCAGGAATT 513  
Qy 290 -----AAATTCA 297  
Db 514 TCACAAGCTGTGCAACAAGATATGTCCACAGCTGCAGAGACAAAAGAAACAGAAATTTTCA 573  
Qy 298 CAAGCTATGGAACAAGAGGTGCCACAGACAGACAGACAGAAAGCAAGAAATTTTCA 357  
Db 574 CAAGCTATGGAACAAGAGGTGCCACAGACAGACAGACAGAAAGCAAGAAATTTTCA 633  
Qy 358 GCTATGGAACAAGAGGTGCCACAGACAGACAGACAGAAAGCAAGAAATTTGAGGTAC 417  
Db 634 GCTATGGAACAAGAGGTGCCACAGACAGACAGACAGAAAGCAAGAGGTGAGTAC 693  
Qy 418 ACATGGAGCTCAAGAGTCAAGTATGACAAATTTGCTGAGAGAGAGATGATAGTGT 477  
Db 694 ACATGGAGCTCAAGAGTCAAGTATGACAAATTTGCTGAGAGAGAGATGATAGTGT 753  
Qy 478 AGTTTGAAGAACTGTGCTGACCTGGCCGGAATGCAACGTTGGCTGTGCTTTTAT 537

Db 754 AGTTTGAAGAACTGTGCTGACCTGGCCGGAATGCAACGTTGGCTGTGCTTTTAT 813  
Qy 538 TCACAACGGTGGGGCTTCGCGCCCTCCGACCGGTGTTCTGACAGGAAGTACAGAAATTTGGG 597  
Db 814 TCACAACGGTGGGGCTTCGCGCCCTCCGACCGGTGTTCTGACAGGAAGTACAGAAATTTGGG 873  
Qy 598 AAATCGCTCTAGCAGAAAGATCGTGTGCTGGCGACAGAGTGAATCTTACAGGGA 657  
Db 874 AAATCGCTCTAGCAGAAAGATCGTGTGCTGGCGACAGAGTGAATCTTACAGGGA 933  
Qy 658 ATGTTCTCTACGTCCTTCTTCTCCCGTTAAGAGATGCAAGCGGAAGAGAGACAGT 717  
Db 934 ATGTTCTCTACGTCCTTCTTCTCCCGTTAAGAGATGCAAGCGGAAGAGAGACAGT 993  
Qy 718 GTCAACAAGTTCAATCTCAGAGAGATGGCCACATCCCAAGCTCCGGTGAACGGAATCATG 777  
Db 994 GTCAACAAGTTCAATCTCAGAGAGATGGCCACATCCCAAGCTCCGGTGAACGGAATCATG 1053  
Qy 778 TCCGACCAAGAAAGCTGTTGTTTCATCATTTGACATGACCTGGGCTGTGCTC 837  
Db 1054 TCCGACCAAGAAAGCTGTTGTTTCATCATTTGACATGACCTGGGCTGTGCTC 1113  
Qy 838 AACAAATGACAAAGCTCTGCAAAAGCTGGCTGAGAGACAGCTCCGTTCAACCTCAT 897  
Db 1114 AACAAATGACAAAGCTCTGCAAAAGCTGGCTGAGAGACAGCTCCGTTCAACCTCAT 1173  
Qy 898 CGCAGTCTGAGAGAAAGTCTGCTCCCTGAGTCTTCTCGATCGTCAACGTCAGAGAC 957  
Db 1174 CGCAGTCTGAGAGAAAGTCTGCTCCCTGAGTCTTCTCGATCGTCAACGTCAGAGAC 1233  
Qy 958 GTGGGACAGAGAAAGTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTTAGTTAGAGA 1017  
Db 1234 GTGGGACAGAGAAAGTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTTAGTTAGAGA 1293  
Qy 1018 ATCTCCGGGAAACAAAGATTCATCTGCTCTCTTGAAGCGCGGGAATTTGTTAGCATGAG 1077  
Db 1294 ATCTCCGGGAAACAAAGATTCATCTGCTCTCTTGAAGCGCGGGAATTTGTTAGCATGAG 1353  
Qy 1078 ACACAAGGTTGCGTGGATCATGACAACCGTGAAGTGTCTGACCAAGTGCAGAGTCCG 1137  
Db 1354 ACACAAGGTTGCGTGGATCATGACAACCGTGAAGTGTCTGACCAAGTGCAGAGTCCG 1413  
Qy 1138 GCCGTGGGCTCTCTCATCTGAGTGGCCCTGACGCTGACGACGTTGGGGAGAGGCTC 1197  
Db 1414 GCCGTGGGCTCTCTCATCTGAGTGGCCCTGACGCTGACGACGTTGGGGAGAGGCTC 1473  
Qy 1198 GCCCTTTTCAACAAGCTTCAACGCTGACGCGCTTTTGTGTTTCATGACTCAAC 1257  
Db 1474 GCCCTTTTCAACAAGCTTCAACGCTGACGCGCTTTTGTGTTTCATGACTCAAC 1533  
Qy 1258 CTTGAGAGCGGTGTCGCGCGCTGTCTCAATCTGAGAGAAAGATGTCTGAAAGCGCTTC 1317  
Db 1534 CTTGAGAGCGGTGTCGCGCGCTGTCTCAATCTGAGAGAAAGATGTCTGAAAGCGCTTC 1593  
Qy 1318 TGCCGTATGCTGTGAGAGGAGTGTGAATAGAAAGTCAAGTGTGTTGATGTTGAGACCTC 1377  
Db 1594 TGCCGTATGCTGTGAGAGGAGTGTGAATAGAAAGTCAAGTGTGTTGATGTTGAGACCTC 1653  
Qy 1378 ATGTTTCAAGAGCTGGGAGATCTGAGCTCCGTCTGTTTCATGATGAATCTTCTTC 1437  
Db 1654 ATGTTTCAAGAGCTGGGAGATCTGAGCTCCGTCTGTTTCATGATGAATCTTCTTC 1713  
Qy 1438 CCAAGACGCACTGTGAGAGTACTAACAATCTTCCACATCTCCAGGACCTTCGT 1497  
Db 1714 CCAAGACGCACTGTGAGAGTACTAACAATCTTCCACATCTCCAGGACCTTCGT 1773  
Qy 1498 GCCGCTTTGATCTAGTGTGAGAGGCTGGAATTCAGCAGCTCTGCGCTCTGATC 1557  
Db 1774 GCCGCTTTGATCTAGTGTGAGAGGCTGGAATTCAGCAGCTCTGCGCTCTGATC 1833  
Qy 1558 GTTGAAGAACAAAGAGTTCATGAGCTTAAACAGGCAAGCTTCAATATCACTCGCTT 1617

Db 1834 GTTGAGAGACAAAGAGTCCATGAGCTTAAACAGGCGAGCTTCATCCACTGCTT 1893  
Qy 1618 TGGATGAAGCGTTCTGTTGGCCCTGAGAGGAAGATAGAGGCGCATCGAGAGTC 1677  
Db 1894 TGGATGAAGCGTTCTGTTGGCCCTGAGAGGAAGATAGAGGCGCATCGAGAGTC 1953  
Qy 1678 CTGCTGGGCTGTCCCGTTCCTCGGGGGTGAAGCAGAAAGCTTCGCACTGGCTCTCTG 1737  
Db 1954 CTGCTGGGCTGTCCCGTTCCTCGGGGGTGAAGCAGAAAGCTTCGCACTGGCTCTCTG 2013  
Qy 1738 TTGGGCTCAGCAGCCTTAATGCCAACCCAGAGAGACCCCTGAGAGCTTCACCTGTCTT 1797  
Db 2014 TTGGGCTCAGCAGCCTTAATGCCAACCCAGAGAGACCCCTGAGAGCTTCACCTGTCTT 2073  
Qy 1798 TTGGGCTCAGCAGCCTTAATGCCAACCCAGAGAGACCCCTGAGAGCTTCACCTGTCTT 1857  
Db 2074 TTGGGCTCAGCAGCCTTAATGCCAACCCAGAGAGACCCCTGAGAGCTTCACCTGTCTT 2133  
Qy 1858 CTTCGAGTTAACAGAACCTTGACCTTGATGACCTTCCTTCGCTCCAGCAGCTGTCCG 1917  
Db 2134 CTTCGAGTTAACAGAACCTTGACCTTGATGACCTTCCTTCGCTCCAGCAGCTGTCCG 2193  
Qy 1918 TATTTGCGGAAAATTCGGGGTGAATGCAAAAGGATCTTCCAAAGATGAGTCCGCTGAG 1977  
Db 2194 TATTTGCGGAAAATTCGGGGTGAATGCAAAAGGATCTTCCAAAGATGAGTCCGCTGAG 2253  
Qy 1978 GATATTCCTGTTGCTCCCTCTATGAGATGCGGGATTAAGACCTTCATTAAGAGAGAGTGGAA 2037  
Db 2254 GATATTCCTGTTGCTCCCTCTATGAGATGCGGGATTAAGACCTTCATTAAGAGAGAGTGGAA 2313  
Qy 2038 GATTTCTGCTCATGCTTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2097  
Db 2314 GATTTCTGCTCATGCTTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2373  
Qy 2098 ATCTGAGCAGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGATCCCACTGCAAG 2157  
Db 2374 ATCTGAGCAGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGATCCCACTGCAAG 2433  
Qy 2158 ATACGAGACCTTGATGTTTGAAGATGCAAGATTAACCTTCGTGTGTGAGAGAGTCTGAGAG 2217  
Db 2434 ATACGAGACCTTGATGTTTGAAGATGCAAGATTAACCTTCGTGTGTGAGAGAGTCTGAGAG 2493  
Qy 2218 ATGCTCATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2277  
Db 2494 ATGCTCATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2553  
Qy 2278 GAGGATGTAAGAGATGCGGTGAGAGCCTTAAACACCCAAATGTTTGTGAGTCTTGG 2337  
Db 2554 GAGGATGTAAGAGATGCGGTGAGAGCCTTAAACACCCAAATGTTTGTGAGTCTTGG 2613  
Qy 2338 AGGCTGATGTTGCTGAGATGAACCAATGCTTTAAGCTGAGATCTCCAAATCTTACG 2397  
Db 2614 AGGCTGATGTTGCTGAGATGAACCAATGCTTTAAGCTGAGATCTCCAAATCTTACG 2673  
Qy 2398 ACCCTCCCGAGCCTGAATCTCTGAGCCTGAGAGAAACAGGTGACAGACCAAGGAGATA 2457  
Db 2674 ACCCTCCCGAGCCTGAATCTCTGAGCCTGAGAGAAACAGGTGACAGACCAAGGAGATA 2733  
Qy 2458 ATGCTCTCATGATGCTTGAAGAGTCTCCAGTGCAGCCTGACAGAGCTGATCTGAG 2517  
Db 2734 AGGCTCTCATGATGCTTGAAGAGTCTCCAGTGCAGCCTGACAGAGCTGATCTGAG 2793  
Qy 2518 GACTGTGAGATCAAGCCAGCGGTTGCCAGAGTCTGGCCTCAGCCCTGTGACAGACCGG 2577  
Db 2794 GACTGTGAGATCAAGCCAGCGGTTGCCAGAGTCTGGCCTCAGCCCTGTGACAGACCGG 2853  
Qy 2578 AGCTTGACACACCTTGCTATCAACACAGCCTGGGGAGAGAGGTGAATCTTACG 2637  
Db 2854 AGCTTGACACACCTTGCTATCAACACAGCCTGGGGAGAGAGGTGAATCTTACG 2913  
Qy 2638 TGTGATCATAGAGCTTCCCACTGTAGTGTGAGAGAGGCTGATGCTGATCAGTGCAC 2697  
Db 2914 TGTGATCATAGAGCTTCCCACTGTAGTGTGAGAGAGGCTGATGCTGATCAGTGCAC 2973

Qy 2698 CTGACAGGCTGCTGTGTTTCTTGCACTTGCGCTTAAGGGTAACTAGCTGACG 2757  
Db 2974 CTGACAGGCTGCTGTGTTTCTTGCACTTGCGCTTAAGGGTAACTAGCTGACG 3033  
Qy 2758 CACTGAGCCTTATGAGTAAACCTGTGAGAGCAATGCGTGAACCTTCGTGAGAGTC 2817  
Db 3034 CACTGAGCCTTATGAGTAAACCTGTGAGAGCAATGCGTGAACCTTCGTGAGAGTC 3093  
Qy 2818 ATGAGAGAACCATCTTGATCTCCAGAGCCTGAGATTGTAAGTGTATCTACCGCC 2877  
Db 3094 ATGAGAGAACCATCTTGATCTCCAGAGCCTGAGATTGTAAGTGTATCTACCGCC 3153  
Qy 2878 GGTGCTGTGAGAGTCTGCTGTGTGATCTCGAGAGACACACCTGAAGAGCTGTAT 2937  
Db 3154 GGTGCTGTGAGAGTCTGCTGTGTGATCTCGAGAGACACACCTGAAGAGCTGTAT 3213  
Qy 2938 CTACAGCAATAGCCTGGTGAAGCGGGGTGCTGCGCTGTGAGAGAGCACTGAAGCA 2997  
Db 3214 CTACAGCAATAGCCTGGTGAAGCGGGGTGCTGCGCTGTGAGAGAGCACTGAAGCA 3273  
Qy 2998 AAGAACAGTGTCTGACAGACCTGGGTTGAAGGATGATGATCTTGTGATGCTGT 3057  
Db 3274 AAGAACAGTGTCTGACAGACCTGGGTTGAAGGATGATGATCTTGTGATGCTGT 3333  
Qy 3058 GAGGACCTCTGCTGGCCCTTCCGCAACCGGATCTGACAGCTTAACTGCTGAG 3117  
Db 3334 GAGGACCTCTGCTGGCCCTTCCGCAACCGGATCTGACAGCTTAACTGCTGAG 3393  
Qy 3118 AATACTCAGTCCCAAGATGATGAAGCTGTGCTGCTTGTGCTGCTGCCAGTCT 3177  
Db 3394 AATACTCAGTCCCAAGATGATGAAGCTGTGCTGCTTGTGCTGCCAGTCT 3453  
Qy 3178 AACTTACAGATTAATGAGCTGTGAAATGAGAGTCCCTGCAATTAAGAGAGTGTG 3237  
Db 3454 AACTTACAGATTAATGAGCTGTGAAATGAGAGTCCCTGCAATTAAGAGAGTGTG 3513  
Qy 3238 GAGGAGTGCAGCTACTAAGCCCGAGTCTGTAATGAGAGTGTGATCTTGTAT 3297  
Db 3514 GAGGAGTGCAGCTACTAAGCCCGAGTCTGTAATGAGAGTGTGATCTTGTAT 3573  
Qy 3298 GAAGTACCGACAC 3312  
Db 3574 GAAGTACCGATAC 3588

RESULT 3  
AX427610  
LOCUS  
DEFINITION Sequence 23 from Patent W00232955.  
ACCESSION AX427610  
VERSION AX427610.1 GI:21537730  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
REFERENCE 1  
AUTHORS Nelson,L.M. and Tong,Z.-B.  
TITLE Human gene critical to fertility  
JOURNAL Patent: WO 0232955-A 23 25-APR-2002;  
GOVERNMENT OF THE UNITED STATES (US)  
LOCATION/Qualifiers  
FEATURES  
source  
1..3900  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
1..3603  
/note="unassigned protein product"  
/codon\_start=1  
/protein\_id="CAD35973.1"  
/db\_xref="GI:21537731"  
/translation="MKVAGLELGAALLSAPRALVTLSTGPTCSILPKNPILFPQNL

SSOPCIKMEGDKSLTPSSYGLWCLYELDKKEEPOTFKELLKKKSSESTCS1POFEIE  
NANVECLALILHEYYGASLAWATSIISFENNMLRTLSKADUDKRSSEBPETATMD  
OGPEKREXVPGISOAVOODSATATKIOEISOAMEORGATATKEEEDISEBPEATMD  
TAATREOGHGGDITMDYKSHMTKFAEBEDRBRSPENTADMPMOTLADGPDNRG  
FRPTVLDHGSIGSGLARIVLCMAOGSLYOGMFSYFLEKDMAEKQRPPTLIR  
FISEWMPDSQAPTEIMSRPERLFIIDGFDDLSVANDKTLCKDWAEXKQRPPTLIR  
SLIKRVLPESEFLIVTVADVOTKLSKSVSPRILVRGISEGBRHHLEIRGEBHQ  
KTQGLRAIINRELLDOCOVAVGSLICVALQLODVESVAPENQOILHEDGVALCEG  
OLTRGVVRCLNLEERVLKRFORMAVEGWNKRSVEDGDLMDGKVESEMLKOA  
MNLTPDSCHEEYVTFPHLSIODPCALVYLRLGLEIEPALCPYVYKTRSMELKOA  
GFTHLSLMKRFPLGLVSEDRPLEVTLIGCPVPLGYKOKLHNVSLIGQOPNATTTG  
DTLDAFCLFETQDEPVRILNLSFOEWLPIINONLDLASSFCLQKEXPIRYKRLRVY  
KGIPRDESAACVPEVPLMMDKTLIEBOWDFCSMLGTHLEHLOLDGSSILTERAM  
KITLCAKLRHPCKIOTLMEFRNAOITPGVOHLMRYMANRNLISNLGTHLEEDVPM  
ACBALKHPKCLLESRLDCCGLTHACVYKISLITLTSPLSKLSLAKNGVTDGVTPL  
SDALRVQCALQKILIEDCGITATGCGSIALMANSSTHLSLKNPVEDDGVALLCE  
RSKRLPHCSLORMLNCHLDTAGCESIALMANSSTHLSLKNPVEDDGVALLCE  
VMRPSCHLODELKVLKCHLTAACESLSVLSRSHLKLDTNLALGDGVALCEG  
LKQKNSVLTBLGLAKGILTDCBALSANRHLTSLNLYONNFPKGMKLSAF  
ACPTSNLQIIGLMKQYFVQIRKLLEEVQLKPRVLDGSHWSFEDEDRYWKNK"

## ORIGIN

Query Match 78.5%; Score 3166.2; DB 6; Length 3900;  
Best Local Similarity 96.2%; Pred. No. 0;  
Matches 3304; Conservative 0; Mismatches 8; Indels 123; Gaps 1;

OY 1 ATGGAAGGAGCAATATGCTCACCCTTTCCAGCTACGGGCTGCATGGTGTCTATGAG 60  
DB 154 ATGGAAGGAGCAATATGCTCACCCTTTCCAGCTACGGGCTGCATGGTGTCTATGAG 213  
OY 61 CTAGACAAGGAAGATTTCAAGCATTTCAAGAAATTTCTTAAGAAATCTTCAAGATCG 120  
DB 214 CTAGACAAGGAAGATTTCAAGCATTTCAAGAAATTTCTTAAGAAATCTTCAAGATCG 273  
OY 121 ACCAATGCTCTATTCCACAGTTTGAATTCAGAAATGCGAACGTGGAAATGTCTGCACTC 180  
DB 274 ACCAATGCTCTATTCCACAGTTTGAATTCAGAAATGCGAACGTGGAAATGTCTGCACTC 333  
OY 181 CTCTTGATGATTTATGAGCATCGTGGCTGGGCTAGTCCATTAGCATTTGAA 240  
DB 334 CTCTTGATGATTTATGAGCATCGTGGCTGGGCTAGTCCATTAGCATTTGAA 393  
OY 241 AACATGAACCTGCGAACCTCTCGGAGAAAGGACGGATGACATGAAAA----- 289  
DB 394 AACATGAACCTGCGAACCTCTCGGAGAAAGGACGGATGACATGAAAAATTTCAACA 453  
OY 290 ----- 289  
DB 454 GAAGATCTGGAAGCAAGCATGATGACCAAGGACCAAGCAAGAAAGTCCAGAAAT 513  
OY 290 -----AAATTTCA 297  
DB 514 TCACAACTGTGCAACAAGATATGTCACAGCTGACAGACAAAGAAAGAGAAATTTCA 573  
OY 298 CAAGCTATGGAACAAGAGGTGCGACACAGCAGACAGACAGAAAGCAAGAAATTTCAACA 357  
DB 574 CAAGCTATGGAACAAGAGGTGCGACACAGCAGACAGACAGAAAGCAAGAAATTTCAACA 633  
OY 358 GCTATGGAACAAGAGGTGCGACACAGCAGACAGACAGAAAGCAAGCAATGAGGTGAC 417  
DB 634 GCTATGGAACAAGAGGTGCGACACAGCAGACAGACAGAAAGCAAGCAATGAGGTGAC 693  
OY 418 ACATGGGACTCAAGAGTCAAGTATGACCAAAATTCGCTGAGGAGGAGATGATAGTCTGT 477  
DB 694 ACATGGGACTCAAGAGTCAAGTATGACCAAAATTCGCTGAGGAGGAGATGATAGTCTGT 753  
OY 478 AGTTTGAAGAACATGCTGCTGACCTGCGGAAATGCAAAAGTGGCTGTGCTTTTAT 537  
DB 754 AGTTTGAAGAACATGCTGCTGACCTGCGGAAATGCAAAAGTGGCTGTGCTTTTAT 813  
OY 538 TCAAGCCGGTGGGCTTCCGGCTTCGCAAGGTGTCTTCGACCGAAAGTCAAGAAATTTGG 597  
DB 814 TCAAGCCGGTGGGCTTCCGGCTTCGCAAGGTGTCTTCGACCGAAAGTCAAGAAATTTGG 873

OY 598 AATCGGCTCTAGCCAGAAAGATCGTCTGTGCTGGCGCAAGGTGACCTTACCAAGGA 657  
DB 874 AATCGGCTCTAGCCAGAAAGATCGTCTGTGCTGGCGCAAGGTGACCTTACCAAGGA 933  
OY 658 ATGTTCTCTAAGCTCTTCTCTCTCCGTTAGAGATGCAAGCGGAAGAAAGAGCACT 717  
DB 934 ATGTTCTCTAAGCTCTTCTCTCTCCGTTAGAGATGCAAGCGGAAGAAAGAGCACT 993  
OY 718 GTCAAGATTCATCTCAGGAGGTGCGCAGACTCCAGGCTCCGGTGAAGAGATCATG 777  
DB 994 GTCAAGATTCATCTCAGGAGGTGCGCAGACTCCAGGCTCCGGTGAAGAGATCATG 1053  
OY 778 TCCGAGCAGAAAGGCTGTTGTTTCATATGACGAGTTTCATGACCTGGCTGTGCTC 837  
DB 1054 TCCGAGCAGAAAGGCTGTTGTTTCATATGACGAGTTTCATGACCTGGCTGTGCTC 1113  
OY 838 AACAAATGACAAAGCTTCGAAAAGCTGGGCTGAGAGAGCAAGCTCCGTTCAACCTCAT 897  
DB 1114 AACAAATGACAAAGCTTCGAAAAGCTGGGCTGAGAGAGCAAGCTCCGTTCAACCTCAT 1173  
OY 898 CGCAGTCTGCTGAGGAAGTCTGCTCCGTAAGTCTTCTGATTCGTCACCGTCAAGAC 957  
DB 1174 CGCAGTCTGCTGAGGAAGTCTGCTCCGTAAGTCTTCTGATTCGTCACCGTCAAGAC 1233  
OY 958 GTGGGCAAGAGAGCTCAAGTCAAGAGTGTGCTCCCGTTACCTGTTAGTTAGAGA 1017  
DB 1234 GTGGGCAAGAGAGCTCAAGTCAAGAGTGTGCTCCCGTTACCTGTTAGTTAGAGA 1293  
OY 1018 ATCTCGGGGAAACAAAGATTCATCTGCTCTTGAAGCGCGGAAATGTTGATGACATCA 1077  
DB 1294 ATCTCGGGGAAACAAAGATTCATCTGCTCTTGAAGCGCGGAAATGTTGATGACATCA 1353  
OY 1078 ACAAGAGGTTGCGTGGATTCATGAACAACCGTGAAGTGTCTGACAGAGTGCAGAGTCCC 1137  
DB 1354 ACAAGAGGTTGCGTGGATTCATGAACAACCGTGAAGTGTCTGACAGAGTGCAGAGTCCC 1413  
OY 1138 GCCGTGGGCTCTCATCTGCGTGGGCTGACAGTGCAGAGACGTGTGGGAGAGAGCTC 1197  
DB 1414 GCCGTGGGCTCTCATCTGCGTGGGCTGACAGTGCAGAGACGTGTGGGAGAGAGCTC 1473  
OY 1198 GCCCTTTCAACCAAGCTTCACAGGCTTCGACGCGCTTTGTTGTTTCAATCACTCAC 1257  
DB 1474 GCCCTTTCAACCAAGCTTCACAGGCTTCGACGCGCTTTGTTGTTTCAATCACTCAC 1533  
OY 1258 CCTGAGGCGTGTGCGCGCTGTCTCAATCTGAGGAAAGATGTCTGAGGCGCTTC 1317  
DB 1534 CCTGAGGCGTGTGCGCGCTGTCTCAATCTGAGGAAAGATGTCTGAGGCGCTTC 1593  
OY 1318 TGCCGTATGCTGTGAGAGGAGTGTGAATGAGAAATCAAGTGTGTTGATGAGTCACTC 1377  
DB 1594 TGCCGTATGCTGTGAGAGGAGTGTGAATGAGAAATCAAGTGTGTTGATGAGTCACTC 1653  
OY 1378 ATGTTCAAGGACTCGGAGAGTGTGAGCTTCGCTGCTGTTTCAATGAACATCTTCTC 1437  
DB 1554 ATGTTCAAGGACTCGGAGAGTGTGAGCTTCGCTGCTGTTTCAATGAACATCTTCTC 1713  
OY 1338 CCAAGACGCACTGTGAGGAGTACTACACTTCTTCAACTCACTCCAGAACTTCTGT 1497  
DB 1714 CCAAGACGCACTGTGAGGAGTACTACACTTCTTCAACTCACTCCAGAACTTCTGT 1773  
OY 1498 GCCGCTTTGATCTAGGTTTGAAGAGGCTGGAATTCAGACAGCTCTGCTGCTGTAC 1557  
DB 1774 GCCGCTTTGATCTAGGTTTGAAGAGGCTGGAATTCAGACAGCTCTGCTGCTGTAC 1833  
OY 1558 GTTGAAGAACAAAGAGTTCATGAGCTTAAACAGGCAAGCTTTCATATTCATCTGCTT 1617  
DB 1834 GTTGAAGAACAAAGAGTTCATGAGCTTAAACAGGCAAGCTTTCATATTCATCTGCTT 1893  
OY 1618 TGAATGAAGGTTTCTTGTGCTGTGAGGCAACGTTAAGAGAGGCACTGAGAGTTC 1677  
DB 1894 TGAATGAAGGTTTCTTGTGCTGTGAGGCAACGTTAAGAGAGGCACTGAGAGTTC 1953



```

Qy 1678 CTGCTGGAGCTGTCCTCCCTGCGGGGTGAAGCAGAGCTTCTGCACTGGTCTCTCTG 1737
Db 1954 CTGCTGGAGCTGTCCTCCCTGCGGGGTGAAGCAGAGCTTCTGCACTGGTCTCTCTG 2013
Qy 1738 TTGGGTGAGCAGCTTAATGCGCACACCCAGAGAGACCTTGGAGCGCTTCCACTGTCTT 1797
Db 2014 TTGGGTGAGCAGCTTAATGCGCACACCCAGAGAGACCTTGGAGCGCTTCCACTGTCTT 2073
Qy 1798 TTGAGACTCAAGACAAAGATTGTTGGCTTGGCACTTAACAGCTTCCAGAGAGTGGG 1857
Db 2074 TTGAGACTCAAGACAAAGATTGTTGGCTTGGCACTTAACAGCTTCCAGAGAGTGGG 2133
Qy 1858 CTTCGATTAAACAGAACTTGAATGCACTTCTTCTGCTCCAGCACTGTCCG 1917
Db 2134 CTTCGATTAAACAGAACTTGAATGCACTTCTTCTGCTCCAGCACTGTCCG 2193
Qy 1918 TATTTGGCGAAATTCGGGTGATGTCAAAGGATTTTCCCAAGATGATGTCGCTGAG 1977
Db 2194 TATTTGGCGAAATTCGGGTGATGTCAAAGGATTTTCCCAAGATGATGTCGCTGAG 2253
Qy 1978 GCATGTCCTGTCCTCTATGATGCGGGATTAAGACCTGATGAGAGAGAGTGGGA 2037
Db 2254 GCATGTCCTGTCCTCTATGATGCGGGATTAAGACCTGATGAGAGAGAGTGGGA 2313
Qy 2038 GATTTCTGCTCATGCTTGGACCAACCACTGCGGAGCTGGACCTGGGAGAGAGC 2097
Db 2314 GATTTCTGCTCATGCTTGGACCAACCACTGCGGAGCTGGACCTGGGAGAGAGC 2373
Qy 2098 ATCTGACAGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGATCCACCTGCAAG 2157
Db 2374 ATCTGACAGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGATCCACCTGCAAG 2433
Qy 2158 ATACAGACCTGATGTTTGAATGACAGATTAACCTCCGTGTGTGACAGACTCTGAGAG 2217
Db 2434 ATACAGACCTGATGTTTGAATGACAGATTAACCTCCGTGTGTGACAGACTCTGAGAG 2493
Qy 2218 ATCTGATGAGCAACCTGAACCTTAAGATCCCTCAACTTGGAGAGGACCCAGCTGAAGGA 2277
Db 2494 ATCTGATGAGCAACCTGAACCTTAAGATCCCTCAACTTGGAGAGGACCCAGCTGAAGGA 2553
Qy 2278 GAGGATGTAAGAGTGGCGTGTGAGACCTTAAACACCCAAATGTTTGTGAGTCTTGG 2337
Db 2554 GAGGATGTAAGAGTGGCGTGTGAGACCTTAAACACCCAAATGTTTGTGAGTCTTGG 2613
Qy 2338 AGGCTGATGCTGTGTGATGACCAAGCTCTTAACTGAGATCTCCAAATCTCTTACG 2397
Db 2614 AGGCTGATGCTGTGTGATGACCAAGCTCTTAACTGAGATCTCCAAATCTCTTACG 2673
Qy 2398 ACCTCCCGAGGCTGAATCTCTGAGCCTGGCAGAGAAACAAAGGTGACAGACAGGAGGA 2457
Db 2674 ACCTCCCGAGGCTGAATCTCTGAGCCTGGCAGAGAAACAAAGGTGACAGACAGGAGGA 2733
Qy 2458 ATGCTCTCAGTATGCTTGAAGTCTCCAGTGCCTGAGAGCTGAGAGCTGATCTGAG 2517
Db 2734 AGGCTCTCAGTATGCTTGAAGTCTCCAGTGCCTGAGAGCTGAGAGCTGATCTGAG 2793
Qy 2518 GACTGTGACATCAACGCAACGGGTGGCCAGAGTCTGGCCCTGAGCCCTGTCAGCAACCG 2577
Db 2794 GACTGTGACATCAACGCAACGGGTGGCCAGAGTCTGGCCCTGAGCCCTGTCAGCAACCG 2853
Qy 2578 AGCTTGAACAACCTGTGCTATCCAAACAGACCTGGGGAACGAGGTGTAATCTTACTG 2637
Db 2854 AGCTTGAACAACCTGTGCTATCCAAACAGACCTGGGGAACGAGGTGTAATCTTACTG 2913
Qy 2638 TGTGATCATAGAGGCTTCCCACTGTAGTCTGCAAGAGGCTGATCTCAATACAGTCCAC 2697
Db 2914 TGTGATCATAGAGGCTTCCCACTGTAGTCTGCAAGAGGCTGATCTCAATACAGTCCAC 2973
Qy 2698 CTGACACAGGCTGCTGTGTGTTTCTTGCACCTTGGCTTAATGAGTAATCATGCTGAG 2757
Db 2974 CTGACACAGGCTGCTGTGTGTTTCTTGCACCTTGGCTTAATGAGTAATCATGCTGAG 3033
Qy 2758 CACGAGACCTTAGATGAACCTGTGTGAAGACATGAGCGTGAAGCTTCTGTGAGAGTC 2817

```

```

Db 3034 CACCTGAGCCTTAGATGAACCTGTGGAAGCAATGGGTGAAGCTTCTGTGCGAGGTC 3093
Qy 2818 ATGAGAGAACCAATCTTGTATCTCCAGACCTGAGAGTGTGTAATGTATCTCACCGCC 2877
Db 3094 ATGAGAGAACCAATCTTGTATCTCCAGACCTGAGAGTGTGTAATGTATCTCACCGCC 3153
Qy 2878 GGTGCTGTGAGAGTCTGTCTGTGTGATCTGAGAGAGCAGACCTGAAGAGCGTGGAT 2937
Db 3154 GGTGCTGTGAGAGTCTGTCTGTGTGATCTGAGAGAGCAGACCTGAAGAGCGTGGAT 3213
Qy 2938 CTCAAGAGCAATGCCCTGGGTGACCGGTGGGTGCTGCGCTGTGAGAGGAGCTGAAGCAA 2997
Db 3214 CTCAAGAGCAATGCCCTGGGTGACCGGTGGGTGCTGCGCTGTGAGAGGAGCTGAAGCAA 3273
Qy 2998 AAGAACAGTGTCTGACAGAGACTCGGCTTGAAGGAGCATGTGACTGATCTGTGCTGT 3057
Db 3274 AAGAACAGTGTCTGACAGAGACTCGGCTTGAAGGAGCATGTGACTGATCTGTGCTGT 3333
Qy 3058 GAGGACCTCTCTGCGCTTCTGCAACCGGCATCTGAACAGCTAAACCTGGTGCAG 3117
Db 3334 GAGGACCTCTCTGCGCTTCTGCAACCGGCATCTGAACAGCTAAACCTGGTGCAG 3393
Qy 3118 AATACTTCAGTCCCAAGAGATGATGAAGCTGTGTGCGCTTGGCTGTCCAGCTCT 3177
Db 3394 AATACTTCAGTCCCAAGAGATGATGAAGCTGTGTGCGCTTGGCTGTCCAGCTCT 3453
Qy 3178 AACTTACAGATTAATGGCTGTGAAATGCGAGTACCTGTGCAATAAGAGAGTGTCTG 3237
Db 3454 AACTTACAGATTAATGGCTGTGAAATGCGAGTACCTGTGCAATAAGAGAGTGTCTG 3513
Qy 3238 GAGGAGTGCAGCTACTCAAGCCCGAGTCCGTAATGACGTAATGGATTTGAT 3297
Db 3514 GAGGAGTGCAGCTACTCAAGCCCGAGTCCGTAATGACGTAATGGATTTGAT 3573
Qy 3298 GAGGATGACCGACAC 3312
Db 3574 GAGGATGACCGGTAC 3588

```

```

RESULT 4
AX704823
LOCUS AX704823 3830 bp DNA linear PAT 04-APR-2003
DEFINITION Sequence 3 from Patent EP1285964.
ACCESSION AX704823
VERSION AX704823.1 GI:29561488
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE 1
AUTHORS Welss, B., Lessl, M., Peters-Kottig, M. and Beckmann, G.
TITLE Human mater proteins
JOURNAL Patent: EP 1285964 A 3 26-FEB-2003;
FEATURES
location/Qualifiers
1..3830
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 77.8%; Score 3139; DB 6; Length 3830;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 3282; Conservative 0; Mismatches 30; Indels 105; Gaps 2;
Qy 1 ATGAGAGAGCAATGCTTCCAGCTTTCAGCTACGGGCTGCAATGCTGTCTATGAG 60
Db 1 ATGAGAGAGCAATGCTTCCAGCTTTCAGCTACGGGCTGCAATGCTGTCTATGAG 60
Qy 61 CTAGACAGAGAAATTTCAACATTCAGAGATTAATTAAGAAATCTTCAGATCG 120

```

Db 61 CTAGACAGAGAAATTTACAGATCTCAAGGAATTACTMAAGAAATCTTCAGAAATCG 120  
Qy 121 ACCACATGCTCTATTTCCACAGTTTGAATCGAGATGCGCAAGTGAATGCTCGACATC 180  
Db 121 ACCACATGCTCTATTTCCACAGTTTGAATCGAGATGCGCAAGTGAATGCTCGACATC 180  
Qy 181 CTTCTGATGATATTTATGAGATCGCTGAGCTGAGCTACGTCATTCATTCAGATCTTTGAA 240  
Db 181 CTTCTGATGATATTTATGAGATCGCTGAGCTGAGCTACGTCATTCATTCAGATCTTTGAA 240  
Qy 241 AACATGAACCTGCGAACCTCTCGGAGAAAGCAGGGATGACATGAAAAAATTTACAA 300  
Db 241 AACATGAACCTGCGAACCTCTCGGAGAAAGCAGGGATGACATGAAAAAATTTACAA 300  
Qy 301 GCT-----ATGAAACAAGAGTGCACAGAGCAGAGACAGAAAGACA- 344  
Db 301 GATCCTGAAGCAAGATGACTGACCAAGACCAAGCAAGAAAGATGTCAGAAATTA 360  
Qy 345 ----- 344  
Db 361 TATGCAATGACTAAGCTTATCTTGGGGGTGCTGACATCTGACCTGCAATATTAACAC 420  
Qy 345 -----AATAATTTCAAGCTATGGAACAAGAGT 375  
Db 421 AAGTATGTTGAATTCATTTCTTTTGCAGAAATTTCAAGCTATGGAACAAGAGT 480  
Qy 376 GCCACAGCAGAGAGACAGAAACAAGACATGAGGTGACATGAGGACTTAACAAGT 435  
Db 481 GCCACAGCAGAGAGACAGAAACAAGACATGAGGTGACATGAGGACTTAACAAGT 540  
Qy 436 CACGTGATGACCAATTCGTGAGAGAGAGATGACGTGATGTTTGAACAACATGCT 495  
Db 436 CACGTGATGACCAATTCGTGAGAGAGAGATGACGTGATGTTTGAACAACATGCT 495  
Qy 541 CACGTGATGACCAATTCGTGAGAGAGAGATGACGTGATGTTTGAACAACATGCT 600  
Db 496 GCTGATCGGCGGAAATGCAAACTTGCTGCTGCTTTGATTCAGACCGGTGGGCTTC 555  
Qy 601 GCTGATCGGCGGAAATGCAAACTTGCTGCTGCTTTGATTCAGACCGGTGGGCTTC 660  
Db 556 CGGCTCGCAGGAGTCTGCAACGAAAGTCAGGAATTTGGGAAATCGGCTCTAGCCGA 615  
Qy 661 CGGCTCGCAGGAGTCTGCAACGAAAGTCAGGAATTTGGGAAATCGGCTCTAGCCGA 720  
Db 616 AAGATCGTGTGCTGCGCGCAAGTGAAGTCTTACAGGAAATGTTCTCTCAAGTCTTC 675  
Qy 721 AAGATCGTGTGCTGCGCGCAAGTGAAGTCTTACAGGAAATGTTCTCTCAAGTCTTC 780  
Db 676 TTCTCTCCCGTTAAGAGATGACGCGGAAAGAGAGACAGTGTCAAGATTCATCTCC 735  
Qy 781 TTCTCTCCCGTTAAGAGATGACGCGGAAAGAGAGAGTGTCAAGATTCATCTCC 840  
Db 736 AAGGAGTGGCAGACTCCAGGCTCCGCTGACGAGATCATGTCCCGACAGAAAGGCTG 795  
Qy 841 AAGGAGTGGCAGACTCCAGGCTCCGCTGACGAGATCATGTCCCGACAGAAAGGCTG 900  
Db 796 TTGTTTCATATGACGTTTGCATGACCTGAGCTCTGCTCAACAATGACACAAAGCTC 855  
Qy 901 TTGTTTCATATGACGTTTGCATGACCTGAGCTCTGCTCAACAATGACACAAAGCTC 960  
Db 856 TGCAGAAAGCTGGGCTGAGAGAGAGCTCTGTTCAACCTTCATACGAGTCTGTGAGAG 915  
Qy 961 TGCAGAAAGCTGGGCTGAGAGAGAGCTCTGTTCAACCTTCATACGAGTCTGTGAGAG 1020  
Db 916 GTCCGAGCTCCGAGTCTTCTGTGATGTCACCGTCAAGAGAGTGGGAGCAGAGAGCTC 975  
Qy 1021 GTCCGAGCTCCGAGTCTTCTGTGATGTCACCGTCAAGAGAGTGGGAGCAGAGAGCTC 1080  
Db 976 AAGTCAAGAGTCTGTTCTCCCGTTACCTGTTAAGTGAAGAAATCTCCGGGAAACAAAGA 1035  
Qy 1081 AAGTCAAGAGTCTGTTCTCCCGTTACCTGTTAAGTGAAGAAATCTCCGGGAAACAAAGA 1140  
Db 1036 ATTCACATGCTCTTGAAGCGGGGATTTGTGAGATCAAGAAACAAGAGGTTGCTGCG 1095  
Qy 1141 ATTCACATGCTCTTGAAGCGGGGATTTGTGAGATCAAGAAACAAGAGGTTGCTGCG 1200

Qy 1096 ATCATGAACAACGCTGAGTGTCTGACAGATGCGCAGGTGCGCGGCTCTTCATC 1155  
Db 1201 ATCATGAACAACGCTGAGTGTCTGACAGATGCGCAGGTGCGCGGCTCTTCATC 1260  
Qy 1156 TGCCTGCGCTCGACGCTGACAGAGAGTGTGGGGAGAGGCTCGCCCTTCAACAAACG 1215  
Db 1261 TGCCTGCGCTCGACGCTGACAGAGAGTGTGGGGAGAGGCTCGCCCTTCAACAAACG 1320  
Qy 1216 CTCACAGGCGTCAACGCGCTTTTGTGTTCATCAGTCAACCCCTCGAGGCGTGTCCG 1275  
Db 1321 CTCACAGGCGTCAACGCGCTTTTGTGTTCATCAGTCAACCCCTCGAGGCGTGTCCG 1380  
Qy 1276 CGCTGTCTCAATCTGAGAGAAAGTGTCTGAAAGCGCTTTCGCGTATGCTGTGAG 1335  
Db 1381 CGCTGTCTCAATCTGAGAGAAAGTGTCTGAAAGCGCTTTCGCGTATGCTGTGAG 1440  
Qy 1336 GAGTGTGAAATAGAGAGTCAAGTGTGATGTGACGACTCATGTTCAAGACTCGG 1395  
Db 1441 GAGTGTGAAATAGAGAGTCAAGTGTGATGTGACGACTCATGTTCAAGACTCGG 1500  
Qy 1396 GAGTGTGAGCTCGTGTCTGTTTCAATGAACAATCCTTCCACAGCAGCTGTGAG 1455  
Db 1501 GAGTGTGAGCTCGTGTCTGTTTCAATGAACAATCCTTCCACAGCAGCTGTGAG 1560  
Qy 1456 GAGTACTACACCTTCTTCCACTCAGCTTCCAGGACTTCTGTGCGGCTTGTACTACGTG 1515  
Db 1561 GAGTACTACACCTTCTTCCACTCAGCTTCCAGGACTTCTGTGCGGCTTGTACTACGTG 1620  
Qy 1516 TTGAGGCGCTCGAAATGACAGCCAGCTCTGCGCTTGTACGTGGAAGACAAAGAG 1575  
Db 1621 TTGAGGCGCTCGAAATGACAGCCAGCTCTGCGCTTGTACGTGGAAGACAAAGAG 1680  
Qy 1576 TCCATGAGGCTTAAACAGGAGGCTTCATATCCACTGCTTGGATGAAGCGTTCTTG 1635  
Db 1681 TCCATGAGGCTTAAACAGGAGGCTTCATATCCACTGCTTGGATGAAGCGTTCTTG 1740  
Qy 1636 TTTGGCTCTGTGAGCGAAGAGTAAAGAGGCACTGAGAGTCTGCTGAGCTGCTCCGTT 1695  
Db 1741 TTTGGCTCTGTGAGCGAAGAGTAAAGAGGCACTGAGAGTCTGCTGAGCTGCTCCGTT 1800  
Qy 1696 CCCCTGGGGGTGAAGACAAAGCTTCTGCACTGAGTCTCTGTTGGGTCAAGACCTTAAT 1755  
Db 1801 CCCCTGGGGGTGAAGACAAAGCTTCTGCACTGAGTCTCTGTTGGGTCAAGACCTTAAT 1860  
Qy 1756 GCCACCAACCCAGAGAGACCTGAGAGCGCTTCCACGTCTTTCGAGACTCAAGACAA 1815  
Db 1861 GCCACCAACCCAGAGAGACCTGAGAGCGCTTCCACGTCTTTCGAGACTCAAGACAA 1920  
Qy 1816 GAGTTTGTGCTTGGATTTAAACAGCTTCCAGAAAGTGTGAGTCCGATTAACAGAAC 1875  
Db 1921 GAGTTTGTGCTTGGATTTAAACAGCTTCCAGAAAGTGTGAGTCCGATTAACAGAAC 1980  
Qy 1876 CTGACATTGATGATCTTCTCTGCTTCCAGACACTGTCGCTATTTGCGGAAATTCGG 1935  
Db 1981 CTGACATTGATGATCTTCTCTGCTTCCAGACACTGTCGCTATTTGCGGAAATTCGG 2040  
Qy 1936 GTGAGTGTCAAGAGGATCTTCCAGAGATGAGTCCGCTGAGGAGTGCCTGTGCTCCT 1995  
Db 2041 GTGAGTGTCAAGAGGATCTTCCAGAGATGAGTCCGCTGAGGAGTGCCTGTGCTCCT 2100  
Qy 1996 CTATGATGCGGAGTAAAGACCTTCATTTGAGAGAGAGTGGGAAATTTCTGCTCATGCTT 2055  
Db 2101 CTATGATGCGGAGTAAAGACCTTCATTTGAGAGAGAGTGGGAAATTTCTGCTCATGCTT 2160  
Qy 2056 GGCACCAACCCACACTGCTGCGCAGCTGACCTGGGACAGACATCTGACAGAGCGGCGC 2115  
Db 2161 GGCACCAACCCACACTGCTGCGCAGCTGACCTGGGACAGACATCTGACAGAGCGGCGC 2220  
Qy 2116 ATGAAGACCTGTGTGCAAGCTGAGGAGATCCCACTGCAAGATACAGACCTGTGATTT 2175  
Db 2221 ATGAAGACCTGTGTGCAAGCTGAGGAGATCCCACTGCAAGATACAGACCTGTGATTT 2280

2176 AGAATGACAGATTACCCCTGATGACAGACCTCTGAGAGATGTCATGCGCAACCT 2235  
2281 AGAATGACAGATTACCCCTGATGACAGACCTCTGAGAGATGTCATGCGCAACCT 2340  
2236 AACCTAAGATCCCTCACTTGGAGGACCCACTGAGAGAGAGATGTAAGATGCG 2295  
2341 AACCTAAGATCCCTCACTTGGAGGACCCACTGAGAGAGAGATGTAAGATGCG 2400  
2236 TGTGAAGCTTAAACACCCAAATGTTGTGAGAGCTTTAGAGCTGATGCTGGA 2355  
2401 TGTGAAGCTTAAACACCCAAATGTTGTGAGAGCTTTAGAGCTGATGCTGGA 2460  
2356 TTGACCCATGCTGTTACCTGAAGATCTCCAAATCTCCAGACCTCCCGCAGCTGAA 2415  
2461 TTGACCCATGCTGTTACCTGAAGATCTCCAAATCTCCAGACCTCCCGCAGCTGAA 2520  
2416 TCTCTGAGCTTGGAGAGAAACAGAGTGAACAGACAGAGATGATGCTCTCAGTATGC 2475  
2521 TCTCTGAGCTTGGAGAGAAACAGAGTGAACAGACAGAGATGATGCTCTCAGTATGC 2580  
2476 TTGAAGATCTCCCAATGCGCTCTGAGAGCTGATGATGAGAGCTGATGAGCTC 2535  
2581 TTGAAGATCTCCCAATGCGCTCTGAGAGCTGATGATGAGAGCTGATGAGCTC 2640  
2536 ACGGCTTGGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2595  
2641 ACGGCTTGGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2700  
2596 CTATCAACACAGCTGAGGAGAGAGAGTGAATCTACTGCTGATGATGATGAGCTT 2655  
2701 CTATCAACACAGCTGAGGAGAGAGAGTGAATCTACTGCTGATGATGATGAGCTT 2760  
2656 CCCCACTGATCTGAGAGAGCTGATGCTGATGATGATGATGATGATGATGATG 2715  
2761 CCCCACTGATCTGAGAGAGCTGATGCTGATGATGATGATGATGATGATGATG 2820  
2716 GATTTCTTGAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2775  
2821 GATTTCTTGAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2880  
2776 AACCTGAGAGAGAGAGAGCTGAGAGCTTCTGCTGAGAGTATGAGAGACATCTTGT 2835  
2881 AACCTGAGAGAGAGAGAGCTGAGAGCTTCTGCTGAGAGTATGAGAGACATCTTGT 2940  
2836 CATCTCAAGAGCTGAGAGTGTGAAGTGTCACTCAGCGCGCTGCTGAGAGTCTG 2895  
2941 CATCTCAAGAGCTGAGAGTGTGAAGTGTCACTCAGCGCGCTGCTGAGAGTCTG 3000  
2896 TCTCTGATGATCTGAGAGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 2955  
3001 TCTCTGATGATCTGAGAGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 3060  
2956 GGTGACGAGTGGGCTGCTGAGCTGAGAGAGAGAGCTGAGAGAGAGAGAGAGTCTGAG 3015  
3061 GGTGACGAGTGGGCTGCTGAGCTGAGAGAGAGAGCTGAGAGAGAGAGAGAGTCTGAG 3120  
3016 AGACTCGAGTGAAGAGAGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGTCTGAG 3075  
3121 AGACTCGAGTGAAGAGAGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGTCTGAG 3180  
3076 CTCTCTGAG 3135  
3181 CTCTCTGAG 3240  
3136 GGAATGAGAGAGCTGAGTGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 3195  
3241 GGAATGAGAGAGCTGAGTGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 3300  
3196 CTGTGAG 3255  
3301 CTGTGAG 3360  
3256 AAGCCCGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3312

Db 3361 AAGCCCGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3417  
RESULT 5  
LOCUS AX478549 3489 bp DNA linear PAT 12-AUG-2002  
DEFINITION Sequence 3 from Patent WO0248362.  
ACCESSION AX478549  
VERSION AX478549.1 GI:22217318  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Rankumar, J. and Arvizu, C.  
Embryogenesis associated proteins  
Patent: WO 0248362-A 3 20-JUN-2002;  
JOURNAL Incyte Genomics, Inc. (US)  
FEATURES  
source  
1. 3489  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 7474830CB1."  
ORIGIN  
Query Match 77.7%; Score 3133.6; DB 6; Length 3489;  
Best Local Similarity 95.2%; Pred. No. 0;  
Matches 3308; Conservative 0; Mismatches 4; Indels 162; Gaps 1;  
1 ATGAG 60  
1 ATGAG 60  
61 CTAG 120  
61 CTAG 120  
121 ACCAGATGCTATTTCCACAGATTGAAATCGAGATGCAAGTGAATGCTGCACTC 180  
121 ACCAGATGCTATTTCCACAGATTGAAATCGAGATGCAAGTGAATGCTGCACTC 180  
181 CTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
181 CTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
241 AACATGAACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 289  
241 AACATGAACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
290 ----- 289  
301 GATCTGAG 360  
290 ----- 289  
361 TATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
290 ----- 289  
421 AAGTATGTTGAATTCATTTCTTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
319 GCCACAG 378  
481 GCCACAG 540  
379 ACACAG 438  
541 ACACAG 600  
439 GTGATGACCAAAATTCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 498

Db 601 GTGATGACCAATTCGTGAGAGAGAGATGATCGTGAATTTTGAACAACCTGCTCT 660  
Qy 499 GACTGGCCGGAAATTCACAAACGTTGGCTGTGTCTTTATTCAGACCGGTGGGCTTCCGG 558  
Db 661 GACTGGCCGGAAATTCACAAACGTTGGCTGTGTCTTTATTCAGACCGGTGGGCTTCCGG 720  
Qy 559 CCTGCGACGGTGTCTGTGACCGAAAGTCAGGAATTTGGGAAATCGGCTCTAGCCGAGAG 618  
Db 721 CCTGCGACGGTGTCTGTGACCGAAAGTCAGGAATTTGGGAAATCGGCTCTAGCCGAGAG 780  
Qy 619 ATCTGCTGTGTGGGCGCAAGTGAAGTCTTACAGGGAATGTTCTCTAGCTTCTTC 678  
Db 781 ATCTGCTGTGTGGGCGCAAGTGAAGTCTTACAGGGAATGTTCTCTAGCTTCTTC 840  
Qy 679 CTCCCGCTTAAAGAGATGCAACCGGAAAGAGAGACATGTTCACAGATTTATCTCCAG 738  
Db 841 CTCCCGCTTAAAGAGATGCAACCGGAAAGAGAGACATGTTCACAGATTTATCTCCAG 900  
Qy 739 GAGTGGCCAGACTCCAGGCTCCGGTGAACGAGATCATGTCCCGACAGAAAGGCTGTG 798  
Db 901 GAGTGGCCAGACTCCAGGCTCCGGTGAACGAGATCATGTCCCGACAGAAAGGCTGTG 960  
Qy 799 TTCAATCATTTGACGGTTTGCATGACCTTGGGCTCTGTCTCAACAATGACACAAAGCTTCG 858  
Db 961 TTCAATCATTTGACGGTTTGCATGACCTTGGGCTCTGTCTCAACAATGACACAAAGCTTCG 1020  
Qy 859 AAAAGATGGGCTGAGAGACAGCTCCGTTACCTCATACGACAGTCTGTGAGAGAGTTC 918  
Db 1021 AAAAGATGGGCTGAGAGACAGCTCCGTTACCTCATACGACAGTCTGTGAGAGAGTTC 1080  
Qy 919 CTGCTCCCTGAGTCTTCTGTATGTCACCGTCAGAGAGCTGGGACAGAGAGTCCAG 978  
Db 1081 CTGCTCCCTGAGTCTTCTGTATGTCACCGTCAGAGAGCTGGGACAGAGAGTCCAG 1140  
Qy 979 TCAAGAGTGTGTCTCCCGTTACCTGTATGAGGAATCTCCGGGAAACAAAGATTC 1038  
Db 1141 TCAAGAGTGTGTCTCCCGTTACCTGTATGAGGAATCTCCGGGAAACAAAGATTC 1200  
Qy 1039 CACTTGTCTCTTGAAGCGCGGATTTGTGATCATCAAGACACAAAGGTTGCTGCCATC 1098  
Db 1201 CACTTGTCTCTTGAAGCGCGGATTTGTGATCATCAAGACACAAAGGTTGCTGCCATC 1260  
Qy 1099 ATGAACAACCGTGAAGCTGTGACACGATGCCAGGCTCCGCGTGGGCTCTCTCATCTGC 1158  
Db 1261 ATGAACAACCGTGAAGCTGTGACACGATGCCAGGCTCCGCGTGGGCTCTCTCATCTGC 1320  
Qy 1159 GTGGCCCTGACGCTGACAGAGCGTGTGGGAGAGACGTCGCCCTTCAACCAACGCTTC 1218  
Db 1321 GTGGCCCTGACGCTGACAGAGCGTGTGGGAGAGACGTCGCCCTTCAACCAACGCTTC 1380  
Qy 1219 ACAGGCTGTGACGCCGCTTTTGTGTATCATCACTCACCCCTCGAGGCTGTGCCGCTC 1278  
Db 1381 ACAGGCTGTGACGCCGCTTTTGTGTATCATCACTCACCCCTCGAGGCTGTGCCGCTC 1440  
Qy 1279 TGTCTCATCTGAGGAGAAAGTGTCTGAAAGCGCTCTGCGGATAGGCTGTGAGGGA 1338  
Db 1441 TGTCTCATCTGAGGAGAAAGTGTCTGAAAGCGCTCTGCGGATAGGCTGTGAGGGA 1500  
Qy 1339 GTGTGAATAGGAATGATGATTTGATGTGACGACATCATAGTTCAGAGACTCGGAGAG 1398  
Db 1501 GTGTGAATAGGAATGATGATTTGATGTGACGACATCATAGTTCAGAGACTCGGAGAG 1560  
Qy 1399 TCTGAGCTCGTGTCTGTTCATCAATGAACATCTTCTCCAGACAGCCACTGTGAGAG 1458  
Db 1561 TCTGAGCTCGTGTCTGTTCATCAATGAACATCTTCTCCAGACAGCCACTGTGAGAG 1620  
Qy 1459 TACTACACCTTCTTCCACTCATGTCCAGAGACTTCTGTGCGGCTTGTACTACGTGTGA 1518  
Db 1621 TACTACACCTTCTTCCACTCATGTCCAGAGACTTCTGTGCGGCTTGTACTACGTGTGA 1680  
Qy 1519 GAGGCTGTGAATGAGACGAGCTCTGCGCTCTGTAGTTGAGAGACAAAGAGTCC 1578  
|||||

Db 1681 GAGGCTGTGAATGAGACGAGCTCTGCGCTCTGTAGTTGAGAGACAAAGAGTCC 1740  
Qy 1579 ATGAGGCTTAAACAGGAGGCTTCCATATCCACTCGCTTTGAGAGACGTTTCTGTGT 1638  
Db 1741 ATGAGGCTTAAACAGGAGGCTTCCATATCCACTCGCTTTGAGAGACGTTTCTGTGT 1800  
Qy 1639 GGCTGTGAGCGAAGACGTAAAGAGGCACTGAGAGTCTGTGAGGCTGTCCGTTCC 1698  
Db 1801 GGCTGTGAGCGAAGACGTAAAGAGGCACTGAGAGTCTGTGAGGCTGTCCGTTCC 1860  
Qy 1699 CTGGGGGTGAAGCAAGACTTCTGCACTGGGTCTCTGTGTGGGTCAAGCACTTAATGCC 1758  
Db 1861 CTGGGGGTGAAGCAAGACTTCTGCACTGGGTCTCTGTGTGGGTCAAGCACTTAATGCC 1920  
Qy 1759 ACCACCCAGAGACACCTGAGACGCTTCCACTGTTTTCGAGCTCAAGACAAAGAG 1818  
Db 1921 ACCACCCAGAGACACCTGAGACGCTTCCACTGTTTTCGAGCTCAAGACAAAGAG 1980  
Qy 1819 TTTGTGCTTGGATTTAAACAGCTTCCAGAAAGTGTGCTTCCGATTTAACAGAACCTG 1878  
Db 1981 TTTGTGCTTGGATTTAAACAGCTTCCAGAAAGTGTGCTTCCGATTTAACAGAACCTG 2040  
Qy 1879 GACTTGAATGATCTTCTCTGTGCTCCAGACATGCTCCGTAATTTGGGAAATTCGGGTG 1938  
Db 2041 GACTTGAATGATCTTCTCTGTGCTCCAGACATGCTCCGTAATTTGGGAAATTCGGGTG 2100  
Qy 1939 GATGTCAAAGGATCTTCCCAAGAGATGAGTCCGCTGAGGATCTCTGTGCTCCCTA 1998  
Db 2101 GATGTCAAAGGATCTTCCCAAGAGATGAGTCCGCTGAGGATCTCTGTGCTCCCTA 2160  
Qy 1999 TGAATGGGAGTAAGACCTTCATTTGAGAGAGTGGGAAATTTCTGCTCATGTTGGC 2058  
Db 2161 TGAATGGGAGTAAGACCTTCATTTGAGAGAGTGGGAAATTTCTGCTCATGTTGGC 2220  
Qy 2059 ACCACCCACACTTGGGCGAGCTGGAACCTGGGACGACATCTTGAACAGCGGCGCATG 2118  
Db 2221 ACCACCCACACTTGGGCGAGCTGGAACCTGGGACGACATCTTGAACAGCGGCGCATG 2280  
Qy 2119 AAGACCTGTGTGCAAGCTGAGGATCCCACTGGCAAGATGACACCTGATGTTTGA 2178  
Db 2281 AAGACCTGTGTGCAAGCTGAGGATCCCACTGGCAAGATGACACCTGATGTTTGA 2340  
Qy 2179 AATGACAGATTAACCTCTGTGTGACACACTCTGAGAAATCTGATGSCCAACCTTAAC 2238  
Db 2341 AATGACAGATTAACCTCTGTGTGACACACTCTGAGAAATCTGATGSCCAACCTTAAC 2400  
Qy 2239 CTAAGATCTCTCAACTTGGAGAGCACTCACCTGAAGAGATGTAAGATGCGTGT 2298  
Db 2401 CTAAGATCTCTCAACTTGGAGAGCACTCACCTGAAGAGATGTAAGATGCGTGT 2460  
Qy 2299 GAAACCTTAAACACCCAAATGTTTGTGAGATCTTTGAGGCTGTGATGCTGTGATTTG 2358  
Db 2461 GAAACCTTAAACACCCAAATGTTTGTGAGATCTTTGAGGCTGTGATGCTGTGATTTG 2520  
Qy 2359 ACCCATGCTGTTCCTGAAGATCTCCCAATCTTTCAGACCTCCCAAGCTGAATCT 2418  
Db 2521 ACCCATGCTGTTCCTGAAGATCTCCCAATCTTTCAGACCTCCCAAGCTGAATCT 2580  
Qy 2419 CTGAGCTGTGAGAGAAACAGGTGACAGACCAAGGAGTAATGCTCTCAAGTATGCTTTG 2478  
Db 2581 CTGAGCTGTGAGAGAAACAGGTGACAGACCAAGGAGTAATGCTCTCAAGTATGCTTTG 2640  
Qy 2479 AGAGTCTTCCAGTGGCCCTGTGACAGAGCTGATTACTGAGAGACTGTGGATTCACGCCACG 2538  
Db 2641 AGAGTCTTCCAGTGGCCCTGTGACAGAGCTGATTACTGAGAGACTGTGGATTCACGCCACG 2700  
Qy 2539 GGTTCAGAGTGTGGCTGACGCTGTGACAGCAACGAGCTTGAACACCTGTGCTTA 2598  
Db 2701 GGTTCAGAGTGTGGCTGACGCTGTGACAGCAACGAGCTTGAACACCTGTGCTTA 2760  
Qy 2599 TCCAAACACGCTGGGAAACGAAAGTGAATCTGATGTGATCATGAGGCTTCCC 2658  
Db 2761 TCCAAACACGCTGGGAAACGAAAGTGAATCTGATGTGATCATGAGGCTTCCC 2820  
|||||

OY	1	ATGGAAGAGACAATGCGTCACTCCCTTTCCAGCTACCGGGCTGGCAATGGGTCCTATATAG	60
Db	1	ATGGAAGAGACAATGCGTCACTCTTTCCAGCTACCGGGCTGGCAATGGGTCCTATATAG	60
OY	61	CTAGACAAGAAATTTACACATCTCAAGAAATTACTTAAGAAAGAAATCTTCAGAAATCG	120
Db	61	CTAGACAAGAAATTTACACATCTCAAGAAATTACTTAAGAAAGAAATCTTCAGAAATCG	120
OY	121	ACCAATGCTCTATTTCCAGAGTTTGAATTCAGAAATGCGCAACGTGGAAATGTCTGGCACTC	180
Db	121	ACCAATGCTCTATTTCCAGAGTTTGAATTCAGAAATGCGCAACGTGGAAATGTCTGGCACTC	180
OY	181	CTCTTGGATGATATTATAGAGCAATCGCTGGCGCTGGGCTAGTCCATTATAGCATCTTTGAA	240
Db	181	CTCTTGGATGATATTATAGAGCAATCGCTGGCGCTGGGCTAGTCCATTATAGCATCTTTGAA	240
OY	241	AACATGAACCTGCGAACCCTCTCGGAGAAGGCACGGATGACATGAAAA-----	289
Db	241	AACATGAACCTGCGAACCCTCTCGGAGAAGGCACGGATGACATGAAAAATTCAACAGAA	300
OY	290	-----	289
Db	301	GATCCTGAAGCAACGATGACTGACCAAGACCAAGCAAGAAAAAGTCAGAAAAATAA	360
OY	290	-----	289
Db	361	TATGGCATGACTAAGCTTATCTTGGGGGTGTCGACATCTGTGACTGAMTAAATAAACAC	420
OY	290	-----AAATTTCAAGCTATGGAACAAAGGT	318
Db	421	AAGTATGTTGAATTCATTTCTTTTTCAGAAATTTACAGACTATAGAACAAAGGT	480
OY	319	GCCAAGCAGCAGACAGACAGAAAGAACAGAAATTTCAACAAGCTATGGAACAAAGAGTGC	378
Db	481	GCCAAGCAGCAGACAGACAGAAAGAACAGAAATTTCAACAAGCTATGGAACAAAGAGTGC	540
OY	379	ACAGACGACAGACAGAAAGAACAGACATGAGAGTGAACAATGGGACTACAAAGTCAAC	438
Db	541	ACAGACGACAGACAGAAAGAACAGACATGAGAGTGAACAATGGGACTACAAAGTCAAC	600
OY	439	GTCATGACCAAAATTCGCTGAGAGAGAGAGATGAAGTGTAGTTTGAAGAACATGCTGT	498
Db	601	GTCATGACCAAAATTCGCTGAGAGAGAGAGATGAAGTGTAGTTTGAAGAACATGCTGT	660
OY	499	GACTGGCCGGAATATGCAAACTTTGGCTGGTCTTTTGATTAAGACCGGTGGGGCTTCCGG	558
Db	661	GACTGGCCGGAATATGCAAACTTTGGCTGGTCTTTTGATTAAGACCGGTGGGGCTTCCGG	720
OY	559	CCTGGCACGGTGTTCTGACGGAAGTCAAGAAATTTGGAAATCGGCTCTACCGAAGG	618
Db	721	CCTGGCACGGTGTTCTGACGGAAGTCAAGAAATTTGGAAATCGGCTCTAGCCAGAAGG	780
OY	619	ATCGTGTCTGTCTGGGCGCAAGGTGACCTTACAGGGAATGTTCTCTACGTCTTCTTC	678
Db	781	ATCGTGTCTGTCTGGGCGCAAGGTGACCTTACAGGGAATGTTCTCTACGTCTTCTTC	840
OY	679	CTCCCGGTTAGAGAGATTCACGCGGAAGAGAGACAGTGTCAAGAGTTCAATCTCCAGG	738
Db	841	CTCCCGGTTAGAGAGATTCACGCGGAAGAGAGACAGTGTCAAGAGTTCAATCTCCAGG	900
OY	739	GAGTGGCCAGACTCCAGGCTCCGGGTGACGAGATCATGTCCGACCAAGAAAGCGTGTG	798
Db	901	GAGTGGCCAGACTCCAGGCTCCGGGTGACGAGATCATGTCCGACCAAGAAAGCGTGTG	960
OY	799	TTTCAATCAATGACGGTTTGCATGACCTGGGCTCTGTCTTCAACAATGACAAAGCTCTGC	858
Db	961	TTTCAATCAATGACGGTTTGCATGACCTGGGCTCTGTCTTCAACAATGACAAAGCTCTGC	1020
OY	859	AAAAGCTGGGCTGAGAGACAGCTCCGGTTCAACCTCATACGACAGCTGCTGAGGAAGGTC	918
Db	1021	AAAAGCTGGGCTGAGAGACAGCTCCGGTTCAACCTCATACGACAGCTGCTGAGGAAGGTC	1080
OY	919	CTGCTCCCTGAGTCTCTTCTGATGCTACCGGTCAAGACGTGGGCAAGAAAGCTCAAG	978

Db 1081 CTGCTCCCTGAGTCTTCTCTGATCGTCAACCGTCAGAGACGTGGGCACAGAGAGCTCAAG 1140  
Qy 979 TCAGAGGTCGTGTCTCCCGCTTAACTGTATGTAAGAAATTCGCGGGAAACAAAGATC 1038  
Db 1141 TCAGAGGTCGTGTCTCCCGCTTAACTGTATGTAAGAAATTCGCGGGAAACAAAGATC 1200  
Qy 1039 CACTTGTCTTGTAGACGCGGAGATTGTGAGCATCAGAAAGACAAGAGGTTGGTSCATC 1098  
Db 1201 CACTTGTCTTGTAGACGCGGAGATTGTGAGCATCAGAAAGACAAGAGGTTGGTSCATC 1260  
Qy 1099 ATGAACAACCGTAGCTCTCTGACCAAGTGCAGAGTGCCTGCGGTGGGCTCTTCATCTGC 1158  
Db 1261 ATGAACAACCGTAGCTCTCTGACCAAGTGCAGAGTGCCTGCGGTGGGCTCTTCATCTGC 1320  
Qy 1159 GTGGCCCTGCAAGCTGCAAGACGTGTGGGGAGAGACGTGCGCCCTTCAACAAACGCTC 1218  
Db 1321 GTGGCCCTGCAAGCTGCAAGACGTGTGGGGAGAGACGTGCGCCCTTCAACAAACGCTC 1380  
Qy 1219 ACAGGCTGCAAGCGGCTTTTGTGTTCATCAGCTCAACCCCTGAGAGGCTGGCCGCGC 1278  
Db 1381 ACAGGCTGCAAGCGGCTTTTGTGTTCATCAGCTCAACCCCTGAGAGGCTGGCCGCGC 1440  
Qy 1279 TGTCTCAATCTGAGAGAAAGATGTCTTGAAGCGCTTCTGCGGTATGCTGTGAGGGA 1338  
Db 1441 TGTCTCAATCTGAGAGAAAGATGTCTTGAAGCGCTTCTGCGGTATGCTGTGAGGGA 1500  
Qy 1339 GTGTGGAATGAGAAATCAAGTGTGTAAGTGTGACGACCTCATGTGTTCAAGAGATCGGGAG 1398  
Db 1501 GTGTGGAATGAGAAATCAAGTGTGTAAGTGTGACGACCTCATGTGTTCAAGAGATCGGGAG 1560  
Qy 1399 TCTGAGCTCCGTGTCTGTTCACATGAAACATCTTCTCCAGACAGCCACTGTGAGAG 1458  
Db 1561 TCTGAGCTCCGTGTCTGTTCACATGAAACATCTTCTCCAGACAGCCACTGTGAGAG 1620  
Qy 1459 TACTACACCTTCTTCAACCTCAAGTCTCAAGACCTTCTGTGCGGCTTGTACTAGTGTGA 1518  
Db 1621 TACTACACCTTCTTCAACCTCAAGTCTCAAGACCTTCTGTGCGGCTTGTACTAGTGTGA 1680  
Qy 1519 GAGGGCCGTGAAATGAGACCAAGCTCTGCGCTGTGACGTGTAAGTAAGAAACAAAGAGTCC 1578  
Db 1681 GAGGGCCGTGAAATGAGACCAAGCTCTGCGCTGTGACGTGTAAGTAAGAAACAAAGAGTCC 1740  
Qy 1579 ATGAGACTTAAACAGGACAGCTTCCATATCCACTGCTTGGATGAAAGCGTTTCTTGT 1638  
Db 1741 ATGAGACTTAAACAGGACAGCTTCCATATCCACTGCTTGGATGAAAGCGTTTCTTGT 1800  
Qy 1639 GAGCTGTGAGGAAAGAGTGAAGAGGCCACTGAGAGTCTGTGAGGCTGTCCGCTCC 1698  
Db 1801 GAGCTGTGAGGAAAGAGTGAAGAGGCCACTGAGAGTCTGTGAGGCTGTCCGCTCC 1860  
Qy 1699 CTGGGGGGTGAAGCAAGACCTTCTGACATGGGTCTCTGTGGGTCAACAGCTTAATGCC 1758  
Db 1861 CTGGGGGGTGAAGCAAGACCTTCTGACATGGGTCTCTGTGGGTCAACAGCTTAATGCC 1920  
Qy 1759 ACCACCCAGAGACACCTTGACGCTTCCACTGTCTTTCAGAGACTCAAGACAAAGAG 1818  
Db 1921 ACCACCCAGAGACACCTTGACGCTTCCACTGTCTTTCAGAGACTCAAGACAAAGAG 1980  
Qy 1819 TTTGTTCGCTTGGCATTTAAACAGCTTCCAGAGAGTGTGGCTTCGATTAACAGAACTG 1878  
Db 1981 TTTGTTCGCTTGGCATTTAAACAGCTTCCAGAGAGTGTGGCTTCGATTAACAGAACTG 2040  
Qy 1879 GACTTGATAGACATCTTCTGTGCTCTGACGATGTCCGTATTTGGGAAATTCGGGGTG 1938  
Db 2041 GACTTGATAGACATCTTCTGTGCTCTGACGATGTCCGTATTTGGGAAATTCGGGGTG 2100  
Qy 1939 GATGTCAAGAGGATTTCCCAAGAGATGTCCGCTGAGCATGTCTGTGGTCCCTCTA 1998  
Db 2101 GATGTCAAGAGGATTTCCCAAGAGATGTCCGCTGAGCATGTCTGTGGTCCCTCTA 2160  
Qy 1999 TGGATGCGGATAGACCTCTCATTTGAGAGACGTGGAAATTTCTGCTCAATGCTTGGC 2058  
|||||

Db 2161 TGGATGCGGATAGACCTCTCATTTGAGAGACGTGGAAATTTCTGCTCAATGCTTGGC 2220  
Qy 2059 ACCACCCAGACCTGCGGACGTGACCTGGGACAGACATCTTCAACAGAGCGGGCCATG 2118  
Db 2221 ACCACCCAGACCTGCGGACGTGACCTGGGACAGACATCTTCAACAGAGCGGGCCATG 2280  
Qy 2119 AAGACCTGTGTGCAAGCTGAGGACATCCACCTGCAAGATACAGACCTGTATTTAGA 2178  
Db 2281 AAGACCTGTGTGCAAGCTGAGGACATCCACCTGCAAGATACAGACCTGTATTTAGA 2340  
Qy 2179 AATGACAGATTAACCTCGTGTGACAGACCTCTGAGAAATCGTCATGSCCAACCTTAAC 2238  
Db 2341 AATGACAGATTAACCTCGTGTGTGACAGACCTCTGAGAAATCGTCATGSCCAACCTTAAC 2400  
Qy 2239 CTAAGATCCCTCAACTTGGGAGGACCACTGAAAGAAAGATGTAAGATGGCGGT 2298  
Db 2401 CTAAGATCCCTCAACTTGGGAGGACCACTGAAAGAAAGATGTAAGATGGCGGT 2460  
Qy 2299 GAAACCTTAAACACCCAAATGTTGTGAGATCTTTGAGGCTGATGCTGTGATYG 2358  
Db 2461 GAAACCTTAAACACCCAAATGTTGTGAGATCTTTGAGGCTGATGCTGTGATYG 2520  
Qy 2359 ACCATGCTGTTCATCTGAAGATCTCCCAATCTTACGACCTCCCAAGCTGAATCT 2418  
Db 2521 ACCATGCTGTTCATCTGAAGATCTCCCAATCTTACGACCTCCCAAGCTGAATCT 2580  
Qy 2419 CTGAGCTGGCAGGAAACAAAGTGAACAGACCAAGGAGATATGCTCTCAATGATGCTTG 2478  
Db 2581 CTGAGCTGGCAGGAAACAAAGTGAACAGACCAAGGAGATATGCTCTCAATGATGCTTG 2640  
Qy 2479 AAGATCTCCAGTGGCGCTGACAGACCTGATACCTGAGACCTGTGACATCACAGCACG 2538  
Db 2641 AAGATCTCCAGTGGCGCTGACAGACCTGATACCTGAGACCTGTGACATCACAGCACG 2700  
Qy 2539 GGTGTCAGAGATCTGGGCTCAAGCCCTGTGACAGAACCGGAGCTTGAACACACTGTGCTA 2598  
Db 2701 GGTGTCAGAGATCTGGGCTCAAGCCCTGTGACAGAACCGGAGCTTGAACACACTGTGCTA 2760  
Qy 2599 TCCAAACAACGCTGGGGAAACGAAGTGTAAATCTACGTGTGATCATAGAGCTTCCC 2658  
Db 2761 TCCAAACAACGCTGGGGAAACGAAGTGTAAATCTACGTGTGATCATAGAGCTTCCC 2820  
Qy 2659 CACTGTAGTGTGACAGAGCTGATGCTGAATCAGTCCACCTGACACAGGCTGTGTGT 2718  
Db 2821 CACTGTAGTGTGACAGAGCTGATGCTGAATCAGTCCACCTGACACAGGCTGTGTGT 2880  
Qy 2719 TTTCTTGAACCTTGGCTTATGGTAACTCATGCTGTACGACCTGAGCTTTAGATGAAC 2778  
Db 2881 TTTCTTGAACCTTGGCTTATGGTAACTCATGCTGTACGACCTGAGCTTTAGATGAAC 2940  
Qy 2779 CCTGTGAGAGCAATGAGGTGAAGCTTCTGTGCGAGGTCAATGAGAACATCTTGTCAAT 2838  
Db 2941 CCTGTGAGAGCAATGAGGTGAAGCTTCTGTGCGAGGTCAATGAGAACATCTTGTCAAT 3000  
Qy 2839 CTCAGAGACCTGAGATTGTAAAGTGCATCTCACCGCGGTGCTGTGAGAGTCTGTCC 2898  
Db 3001 CTCAGAGACCTGAGATTGTAAAGTGCATCTCACCGCGGTGCTGTGAGAGTCTGTCC 3060  
Qy 2899 TGTGTGATCTGAGAGACAGACACTGAAGAGCTGTGATCTCACAGACATGCTCTGGGT 2958  
Db 3061 TGTGTGATCTGAGAGACAGACACTGAAGAGCTGTGATCTCACAGACATGCTCTGGGT 3120  
Qy 2959 GACGATGGGGTGTGTGCTGTGCTGTGAGAGGACTGAAGCAAGATGTTCTGACGAGA 3018  
Db 3121 GACGATGGGGTGTGTGCTGTGCTGTGAGAGGACTGAAGCAAGATGTTCTGACGAGA 3180  
Qy 3019 CTGGGTTGAAGGATGTGACATGCTTGTGATTTGTGATGAGGACCTCTCTTGGCCCTT 3078  
Db 3181 CTGGGTTGAAGGATGTGACATGCTTGTGATTTGTGATGAGGACCTCTCTTGGCCCTT 3240  
Qy 3079 TCTGTCAAACCGGACATGTGACCAAGTCTAAACCTGTGTGCAAAATTAATTCAATGCCAAAGGA 3138  
Db 3241 TCTGTCAAACCGGACATGTGACCAAGTCTAAACCTGTGTGCAAAATTAATTCAATGCCAAAGGA 3300  
|||||





QY	1438	CCAGACACCACTGGAGAGATCAATCAACCTTTTCCACCTCAAGTCCAGAGACTTCGT	1437
Dp	3535	CCAGACACCACTGGAGAGATCAATCAACCTTTTCCACCTCAAGTCCAGAGACTTCGT	3594
QY	1498	GCCGCTTGTAAGTGTGAAGGCGCTGGAAATGAGCAGACTTCCTGCCTCTGAC	1557
Dp	3595	GCCGCTTGTAAGTGTGAAGGCGCTGGAAATGAGCAGACTTCCTGCCTCTGAC	3654
QY	1558	GTTAGAGAACAAAGAGGTCCATGAGAGCTTAAACAGGAGGCTTCATATCACTCGCT	1617
Dp	3655	GTTAGAGAACAAAGAGGTCCATGAGAGCTTAAACAGGAGGCTTCATATCACTCGCT	3714
QY	1618	TGATGAAAGCTTTCTTTGTTGGCTCCGTGAGCGAAGAGCTAAGAGGCGCACTGAGAGTC	1677
Dp	3715	TGATGAAAGCTTTCTTTGTTGGCTCCGTGAGCGAAGAGCTAAGAGGCGCACTGAGAGTC	3774
QY	1678	CTGCTGGGCTGTCCCGTTCCTCCGTGGGGGTGAAGCAAGACTTCTGCACTGGGTCTCTG	1737
Dp	3775	CTGCTGGGCTGTCCCGTTCCTCCGTGGGGGTGAAGCAAGACTTCTGCACTGGGTCTCTG	3834
QY	1738	TTGGGTCAGAGCCTTAATGCCAACCCACAGAGACACCTTGAGAGGCTTCACTGCTCT	1797
Dp	3835	TTGGGTCAGAGCCTTAATGCCAACCCACAGAGACACCTTGAGAGGCTTCACTGCTCT	3894
QY	1798	TTGGAAGCTCAAGACAAAGAGTTTGTCCTTGCCATTAAACAGCTTCAAGAAAGTGTG	1857
Dp	3895	TTGGAAGCTCAAGACAAAGAGTTTGTCCTTGCCATTAAACAGCTTCAAGAAAGTGTG	3954
QY	1858	CTTCGAGTTAACGAAAGCTGACCTGAACCTTAATGATCTTCCTGCTCCAGACCTGTCCG	1917
Dp	3955	CTTCGAGTTAACGAAAGCTGACCTGATGATCTTCCTGCTCCAGACCTGTCCG	4014
QY	1918	TATTTGCGGAAAATTCGGGTGGATGTCAAGGGAGCTTCCAGAGATGAGTCCGCTGAG	1977
Dp	4015	TATTTGCGGAAAATTCGGGTGGATGTCAAGGGAGCTTCCAGAGATGAGTCCGCTGAG	4074
QY	1978	GCAATGCTCTGTGTCTCCCTCTATGATGCGGGATTAAGACCTTCATTGAGAGACGTGGAA	2037
Dp	4075	GCAATGCTCTGTGTCTCCCTCTATGATGCGGGATTAAGACCTTCATTGAGAGACGTGGAA	4134
QY	2038	GATTTCCTGCTCAATGCTTGGGACCAACCAACCTGCGGAGCTGGAACCTGGGACGACG	2097
Dp	4135	GATTTCCTGCTCAATGCTTGGGACCAACCAACCTGCGGAGCTGGAACCTGGGACGACG	4194
QY	2098	ATCCTGACAGAGCGGGCCATGAGACCTGTGTGTCCAAAGCTGAGGCATCCACTGCAAG	2157
Dp	4195	ATCCTGACAGAGCGGGCCATGAGACCTGTGTGTCCAAAGCTGAGGCATCCACTGCAAG	4254
QY	2158	ATACGAGACCCGATGTTTAAAGAAATGCAACAGATTACCCGTGTGTGACACCTCTGAGGA	2217
Dp	4255	ATACGAGACCCGATGTTTAAAGAAATGCAACAGATTACCCGTGTGTGACACCTCTGAGGA	4314
QY	2218	ATGCTGATGAGCAACCGTAAACCTTAAGATCCCTCAACTTGGAGGACCAACCTGAAGAA	2277
Dp	4315	ATGCTGATGAGCAACCGTAAACCTTAAGATCCCTCAACTTGGAGGACCAACCTGAAGAA	4374
QY	2278	GAGGATGTAAAGATGCGTGTGAAGCCTTAAACACCCAAATATGTTTGTGAGACTTTTG	2337
Dp	4375	GAGGATGTAAAGATGCGTGTGAAGCCTTAAACACCCAAATATGTTTGTGAGACTTTTG	4434
QY	2338	AGGCTGAGATTCTGTGATGTGAACCATATGCTGTTTACCTGAAGATCTCCAAATCCTTACG	2397
Dp	4435	AGGCTGAGATTCTGTGATGTGAACCATATGCTGTTTACCTGAAGATCTCCAAATCCTTACG	4494
QY	2398	ACCTGCCCCAGCCTGAAATCTCTGAGCCTGCGAGGAAACAAAGTGAACAGACCAAGGAGTA	2457
Dp	4495	ACCTGCCCCAGCCTGAAATCTCTGAGCCTGCGAGGAAACAAAGTGAACAGACCAAGGAGTA	4554
QY	2458	ATGCTCTCAATGATCCTTGAAGTCTCCAGATGCGGCTCGAGAACCTATACTGGAG	2517
Dp	4555	ATGCTCTCAATGATCCTTGAAGTCTCCAGATGCGGCTCGAGAACCTATACTGGAG	4614

QY	2518	AACTGTGGCAATCAAGGCAAGGGTTGGCAAGAGTCTGGGCTCAGGCCCTGGTACGAAACCGG	2577
Db	4615	GACTGTGGCAATCAAGGCAAGGGTTGGCAAGAGTCTGGGCTCAGGCCCTGGTACGAAACCGG	4674
QY	2578	AGCTTGAACAACCTGTGGCTAATCCAAACAACAGCCTGGGGAAACGAAGGTAAATCTACTG	2637
Db	4675	AGCTTGAACAACCTGTGGCTAATCCAAACAACAGCCTGGGGAAACGAAGGTAAATCTACTG	4734
QY	2638	TGTGTGATCCAGAGAGCTTCCCACTGTAGTCTGAGAGGCGTGAAGCTGAATCAAGTCCAC	2697
Db	4735	TGTGTGATCCAGAGAGCTTCCCACTGTAGTCTGAGAGGCGTGAAGCTGAATCAAGTCCAC	4799
QY	2698	CTGGAACAAGCTGGCTGTGGTTTCTTGCACTTGCCTTAATAGGTAACTCAGCTGACG	2757
Db	4795	CTGGAACAAGCTGGCTGTGGTTTCTTGCACTTGCCTTAATAGGTAACTCAGCTGACG	4854
QY	2758	CACCTGAGCCTTACATGAACCCCTGTGGAAACAATGGCGTAAGCTTCTGTGGAGGTC	2817
Db	4855	CACCTGAGCCTTACATGAACCCCTGTGGAAACAATGGCGTAAGCTTCTGTGGAGGTC	4914
QY	2818	ATGAGAGAACCATCTTGTCACTCCAGAGCCTGAGAGTTGTAAAGTGCATCTCACGCC	2877
Db	4915	ATGAGAGAACCATCTTGTCACTCCAGAGCCTGAGAGTTGTAAAGTGCATCTCACGCC	4974
QY	2878	GCGTGTCTGTAGAGTCTGTCTTGTGTGATCTCGAGAGCAGACAACCTGAAGAGCTTGAT	2937
Db	4975	GCGTGTCTGTAGAGTCTGTCTTGTGTGATCTCGAGAGCAGACAACCTGAAGAGCTTGAT	5034
QY	2938	CTCAGCGACATAGCCCTGGGTGACGAGTGGGGTTCTGGCGCTGTGCAGAGGACTGAAGAA	2997
Db	5035	CTCAGCGACATAGCCCTGGGTGACGAGTGGGGTTCTGGCGCTGTGTGCAGAGGACTGAAGAA	5094
QY	2998	AAGAACAAGTGTCTGACGAGACTCGGGTGTGAAGGCATGTGGACTGACTTGTGATGTCGT	3057
Db	5095	AAGAACAAGTGTCTGACGAGACTCGGGTGTGAAGGCATGTGGACTGACTTGTGATGTCGT	5154
QY	3058	GAGGCACTCTCCTTGGCCCTTCTGTGCAACCGGATCTGAACCAATCTAAACCTGGTGAG	3117
Db	5155	GAGGCACTCTCCTTGGCCCTTCTGTGCAACCGGATCTGAACCAATCTAAACCTGGTGAG	5214
QY	3118	AATAACTTCACTGCCAAGAAGATGAAGCTGTGTTCCGCGCTTGTGCTGTCCACGCTCT	3177
Db	5215	AATAACTTCACTGCCAAGAAGATGAAGCTGTGTTCCGCGCTTGTGCTGTCCACGCTCT	5274
QY	3178	AACTTACAGATAATTGG 3194	
Db	5275	AACTTACAGATAATTGG 5291	

RESULT 8				
AX459873	AX459873	6939 bp	DNA	Linear
LOCUS	Sequence 16 from Patent WO0240668.			
DEFINITION	AX459873			
ACCESSION	AX459873.1			
VERSION	GI:21725645			
KEYWORDS				
SOURCE				
ORGANISM	Homo sapiens (human)			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1			
AUTHORS	Tschopp, J. and Martinon, F.			
TITLE	Proteins and dna sequences underlying these proteins used for treating inflammations			
JOURNAL	Patent: WO 0240668-A 16 23-MAY-2002;			
	Apotech Research and Development Ltd. (CH)			
FEATURES	Location/Qualifiers			
source	1..6939			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
	/note="NALP5/FY8_dna /GENSCAN_predicted_CDS_1/ 6939_bp"			
ORIGIN				





```
Db 5155 GCATGCTCTGTGTCCTCTCTATGATGCGGGATAAGACCTCATTTAGAGACAGTGGAA 5214
Qy 2038 GATTTCTGCTCAGTCTTGAGACCCACCAACCTGCGGAGCTGGACCTGGCGAGC 2097
Db 5215 GATTTCTGCTCAGTCTTGAGACCCACCAACCTGCGGAGCTGGACCTGGCGAGC 5274
Qy 2098 ATCTGACAGAGCGGGCATGAGAACCTGTGTGCAAGCTGAGGACATCCACCTGCAAG 2157
Db 5275 ATCTGACAGAGCGGGCATGAGAACCTGTGTGCAAGCTGAGGACATCCACCTGCAAG 5334
Qy 2158 ATACAGACCTTGATTTTAAAGATGACATTAACCTTGTGTGAGACCTTGTGAGA 2217
Db 5335 ATACAGACCTTGATTTTAAAGATGACATTAACCTTGTGTGAGACCTTGTGAGA 5394
Qy 2218 ATTCGTCAGGCAACCGTAACCTTAAGATCCCTCAACCTTGGAGGACCACTGAAAGAA 2277
Db 5395 ATTCGTCAGGCAACCGTAACCTTAAGATCCCTCAACCTTGGAGGACCACTGAAAGAA 5454
Qy 2278 GAGGATGTAGGATGCGGTGTGAAACCTTAAACACCCAAATGTTTGTGAGTCTTTG 2337
Db 5455 GAGGATGTAGGATGCGGTGTGAAACCTTAAACACCCAAATGTTTGTGAGTCTTTG 5514
Qy 2338 AGGCTGATTTGTGTGATTTGAACCATGCTGTTAACCTGAAGATTCCTCAATCTTAAG 2397
Db 5515 AGGCTGATTTGTGTGATTTGAACCATGCTGTTAACCTGAAGATTCCTCAATCTTAAG 5574
Qy 2398 ACCCTCCCGAGCGTGAATCTCTGAGCGTGGAGGAAACAAAGGTAGACAGACAGGAGATA 2457
Db 5575 ACCCTCCCGAGCGTGAATCTCTGAGCGTGGAGGAAACAAAGGTAGACAGACAGGAGATA 5634
Qy 2458 ATGCTCTCAGTGTAGTCTTGAAGTCTTCCAGTGCAGCGCTGACAGAACCTGATCTGAG 2517
Db 5635 ATGCTCTCAGTGTAGTCTTGAAGTCTTCCAGTGCAGCGCTGACAGAACCTGATCTGAG 5694
Qy 2518 GACTGTGCAATCAAGCCACGGGTTGCCAGAGTCTGCTCTGACCTTGTGACAGACCGG 2577
Db 5695 GACTGTGCAATCAAGCCACGGGTTGCCAGAGTCTGCTCTGACCTTGTGACAGACCGG 5754
Qy 2578 AGCTTGAACAACCTGTGCTATCCAAACAAGCCGGGGGAAACGAAGGTAAATCTACTAG 2637
Db 5755 AGCTTGAACAACCTGTGCTATCCAAACAAGCCGGGGGAAACGAAGGTAAATCTACTAG 5814
Qy 2638 TGTGATCATGAGAGCTTCCCACTGTAGTCTGACAGAGGCTGATCTGAATCAAGTGCAC 2697
Db 5815 TGTGATCATGAGAGCTTCCCACTGTAGTCTGACAGAGGCTGATCTGAATCAAGTGCAC 5874
Qy 2698 CTGACACGCGCTGCTGTGTTTCTTGACCTTGGCTTAAAGGTAATCATGCTGACG 2757
Db 5875 CTGACACGCGCTGCTGTGTTTCTTGACCTTGGCTTAAAGGTAATCATGCTGACG 5934
Qy 2758 CACCTGAGCCTTGAAGAACCTGTGGAAGCAATGCGTGAAGCTTGTGGGAGGTC 2817
Db 5935 CACCTGAGCCTTGAAGAACCTGTGGAAGCAATGCGTGAAGCTTGTGGGAGGTC 5994
Qy 2818 ATGAGAGAACCATTTGTATCTCCAGACCTGAGAGTGTAAAGTATCATCTCACCCG 2877
Db 5995 ATGAGAGAACCATTTGTATCTCCAGACCTGAGAGTGTAAAGTATCATCTCACCCG 6054
Qy 2878 GCGTGTGTGAGAGTCTGTCTGTGTGATCTTGAGAGACAGACCTGAAGAGCTTGAT 2937
Db 6055 GCGTGTGTGAGAGTCTGTCTGTGTGATCTTGAGAGACAGACCTGAAGAGCTTGAT 6114
Qy 2938 CTCACGGAACATGCGCTGGTGAAGCGTGGGGTGTGCGCTGTGGAAGGACCTGAAGAA 2997
Db 6115 CTCACGGAACATGCGCTGGTGAAGCGTGGGGTGTGCGCTGTGGAAGGACCTGAAGAA 6174
Qy 2998 AAGAAACAATGTTTGAACAGACTCGGTTTGAAGCATGTGACCTTGAATGCTGT 3057
Db 6175 AAGAAACAATGTTTGAACAGACTCGGTTTGAAGCATGTGACCTTGAATGCTGT 6234
Qy 3058 GAGGCACTTCTCTTGGCCCTTTCTGCAACCGGACCTGACCAAGTCTAAACCTGTGAG 3117
```

```
Db 6235 GAGGCACTTCTCTTGGCCCTTTCTGCAACCGGACCTGACCAAGTCTAAACCTGTGAG 6294
Qy 3118 AATAACTTCAAGTCCCAAGGAATGATGAGCTGTGTTGCGCTTGTGCTGTCCACGCTCT 3177
Db 6295 AATAACTTCAAGTCCCAAGGAATGATGAGCTGTGTTGCGCTTGTGCTGTCCACGCTCT 6354
Qy 3178 AACTTACAGATTAATTGG 3194
Db 6355 AACTTACAGATTAATTGG 6371

RESULT 9
AX459891
LOCUS Sequence 34 from Patent WO240668.
DEFINITION AX459891
ACCESSION AX459891
VERSION AX459891.1 GI:21725654
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Teschopp, J. and Martinon, F.
TITLES Proteins and dna sequences underlying these proteins used for
creating inflammations
JOURNAL Patent: WO 0240668-A 34 23-MAY-2002;
Apotech Research and Development Ltd. (CH)
FEATURES
SOURCE
1..6939
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="NALP13/Py17.cdna"

ORIGIN
Query Match 75.8%; Score 3057.8; DB 6; Length 6939;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 3192; Conservative 0; Mismatches 2; Indels 123; Gaps 1;

1 ATGAGAGAGACAAATGCTCTCACTTTTCCAGCTACGCGGCTGCAATGCTGTCTATGAG 60
3055 ATGAGAGAGACAAATGCTCTCACTTTTCCAGCTACGCGGCTGCAATGCTGTCTATGAG 3114
61 CTGACAGAGAAATTTGACATTCATTCAGAAATTAATAAGAAATCTTCAGAAATCG 120
3115 CTGACAGAGAAATTTGACATTCATTCAGAAATTAATAAGAAATCTTCAGAAATCG 3174
121 ACCACATGCTCTATTCACAGTTTGAATGAGATGCCAAGCTGAATGTCTGGCACTC 180
3175 ACCACATGCTCTATTCACAGTTTGAATGAGATGCCAAGCTGAATGTCTGGCACTC 3234
181 CTCTTGATGATATTTATGAGCATCGCTGGCTGGCTGTACGTTCATTAAGCATTTTGAA 240
3235 CTCTTGATGATATTTATGAGCATCGCTGGCTGGCTGTACGTTCATTAAGCATTTTGAA 3294
241 AACATGAACCTGCGCAACCTCTCGAGAGAGGACGAGATGACATGAATAA----- 289
3295 AACATGAACCTGCGCAACCTCTCGAGAGAGGACGAGATGACATGAATAAAGCATTTCA 3354
290 ----- 289
3355 GAAGATCCTGAAGCAAGATGACTGACCAAGAACCAAGAAAGAAAGTGCAGAAATT 3414
290 -----AATTTCA 297
3415 TCACAACTGTGCAACAAGATAGTCCACAGCTGACAGACAAAGAAACAAGAAATTTCA 3474
298 CAAGCTATGGAACAAGAGGTGCCACAGACAGAGACAGAGAAACAAGAAATTTCA 357
3475 CAAGCTATGGAACAAGAGGTGCCACAGACAGAGACAGAGAAACAAGAAATTTCA 3534
358 GCTATGGAACAAGAGGTGCCACAGACAGAGACAGAGAAACAAGAAACAATGAGGTGAC 417
```

Db 3535 GGTATGAAACAAAGAGTGCACAGACGAGAGACAAAGAAACAAGACATGAGGTGAC 3594  
Qy 418 ACATGGAGACTACAAAGTCACTGATGACCAAAATTCGCTGAGAGAGAGATGATCCTGT 477  
Db 3595 ACATGGAGACTACAAAGTCACTGATGACCAAAATTCGCTGAGAGAGAGATGATCCTGT 3654  
Qy 478 AGTTTGAAGAACATGCTGCTGACCTGAGCGGAAATGCAAAACGTTGCTGCTTTTGTAT 537  
Db 3655 AGTTTGAAGAACATGCTGCTGACCTGAGCGGAAATGCAAAACGTTGCTGCTTTTGTAT 3714  
Qy 538 TCAGACCGGTGGGGCTTCCGGCTCGACAGGTGTTCTGACAGGAAATCAGGAATGGG 597  
Db 3715 TCAGACCGGTGGGGCTTCCGGCTCGACAGGTGTTCTGACAGGAAATCAGGAATGGG 3774  
Qy 598 AAATGGCTCTAGACCAAGAGATGCTGCTGCTGGGCGCAAGGTGACTTACAGGGA 657  
Db 3775 AAATGGCTCTAGACCAAGAGATGCTGCTGCTGGGCGCAAGGTGACTTACAGGGA 3834  
Qy 658 ATGTTCTCTAGCTCTTCTTCTCCCGTTAGAGAAATGACAGCGGAAAGAGAGACAT 717  
Db 3835 ATGTTCTCTAGCTCTTCTTCTCCCGTTAGAGAAATGACAGCGGAAAGAGAGACAT 3894  
Qy 718 GTCAAGAGATTATCTCAGAGAGTGGCCAGACTCCAGGCTCCGGTGAACGAGATCATG 777  
Db 3895 GTCAAGAGATTATCTCAGAGAGTGGCCAGACTCCAGGCTCCGGTGAACGAGATCATG 3954  
Qy 778 TCCCGACCAAGAAAGCTGTTGTTTCATCATTTAGACGTTTTCATGACCTGAGCTCTGCTC 837  
Db 3955 TCCCGACCAAGAAAGCTGTTGTTTCATCATTTAGACGTTTTCATGACCTGAGCTCTGCTC 4014  
Qy 838 AACATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGACAGCTTCCTTACCTCATTA 897  
Db 4015 AACATGACACAAAGCTCTGCAAAAGACTGGGCTGAGAGACAGCTTCCTTACCTCATTA 4074  
Qy 898 CCGAGCTCTGAGAAAGTCTGCTGCTCCCTGAGTCCCTTCTGATGCTGACCCGTCAGAGAC 957  
Db 4075 CCGAGCTCTGAGAAAGTCTGCTGCTCCCTGAGTCCCTTCTGATGCTGACCCGTCAGAGAC 4134  
Qy 958 GTGGGACACAGAAAGCTCAAGTCAAGAGTCTGCTCCCGTTACCTGTTAGTTAGAGA 1017  
Db 4135 GTGGGACACAGAAAGCTCAAGTCAAGAGTCTGCTCCCGTTACCTGTTAGTTAGAGA 4194  
Qy 1018 ATCTCCGGGAAACAAAGATCCACTTGTCTTGAACGCGGATTTGTGATCATGAAAG 1077  
Db 4195 ATCTCCGGGAAACAAAGATCCACTTGTCTTGAACGCGGATTTGTGATCATGAAAG 4254  
Qy 1078 AACACAGAGGTTGCGGATCATGAAACAACGTTGAGCTGCGAACGAGCCAGGTCGCC 1137  
Db 4255 AACACAGAGGTTGCGGATCATGAAACAACGTTGAGCTGCGAACGAGTCAGGTCGCC 4314  
Qy 1138 GCGGTGGGCTCTCATCTGCGTGGCCCTGACAGCTGACAGAGCTGGTGGGGAGAGCGTC 1197  
Db 4315 GCGGTGGGCTCTCATCTGCGTGGCCCTGACAGCTGACAGAGCTGGTGGGGAGAGCGTC 4374  
Qy 1198 GCGCCCTTCAACCAACGCTCAACAGGCTGCAACGCGCTTTTGTGTTTATCATGACTCAC 1257  
Db 4375 GCGCCCTTCAACCAACGCTCAACAGGCTGCAACGCGCTTTTGTGTTTATCATGACTCAC 4434  
Qy 1258 CCTGAGAGGTTGCTGGGCTGCTCATCTGAGAGAAAGATTCTGTAAGGCTTTC 1317  
Db 4435 CCTGAGAGGTTGCTGGGCTGCTCATCTGAGAGAAAGATTCTGTAAGGCTTTC 4494  
Qy 1318 TGCCTATGCTGTGAGAGAGTGTGAATAGAAAGTCACTGTTGATGTGACGACTC 1377  
Db 4495 TGCCTATGCTGTGAGAGAGTGTGAATAGAAAGTCACTGTTGATGTGACGACTC 4554  
Qy 1378 ATGCTTCAAGAGACTGGGGAGTGTGAGCTCGCTCTGTTTCAATGAACATCTTCTC 1437  
Db 4555 ATGCTTCAAGAGACTGGGGAGTGTGAGCTCGCTCTGTTTCAATGAACATCTTCTC 4614  
Qy 1438 CAGACAGGCACTGTGAGAGTACTACACTTCTTCCAGCTGCTCAGAGACTTCTGT 1497

Db 4615 CCAGACAGCACTGTGAGAGTACTACACTTCTTCCAGCTGCTCAGAGACTTCTGT 4674  
Qy 1498 GCGCCCTTGTACTAGTGTTAAGGGCCTTGGAAATTCAGCCAGGCTCTGCGCTCTGTAC 1557  
Db 4675 GCGCCCTTGTACTAGTGTTAAGGGCCTTGGAAATTCAGCCAGGCTCTGCGCTCTGTAC 4734  
Qy 1558 GTTGAAGAAACAAGAGGCTCCATGAGAGCTTAAACAGGAGCTTCATATCCACTCGCT 1617  
Db 4735 GTTGAAGAAACAAGAGGCTCCATGAGAGCTTAAACAGGAGCTTCATATCCACTCGCT 4794  
Qy 1618 TGAATGAACCGTTTCTTGTGCTGCTGAGCGAAGAGTAAGAGGCACTGAGAGTC 1677  
Db 4795 TGAATGAACCGTTTCTTGTGCTGCTGAGCGAAGAGTAAGAGGCACTGAGAGTC 4854  
Qy 1678 CTGCTGGGCTGCTCCGTTCCCTGGGGGTGAAGCAAGCTTCTGACTGGGTCTCTG 1737  
Db 4855 CTGCTGGGCTGCTCCGTTCCCTGGGGGTGAAGCAAGCTTCTGACTGGGTCTCTG 4914  
Qy 1738 TTTGGTCAAGACCTTAATGCAACCAACGAGGACACCTGGAAGCCTTCACTGTCTT 1797  
Db 4915 TTTGGTCAAGACCTTAATGCAACCAACGAGGACACCTGGAAGCCTTCACTGTCTT 4974  
Qy 1798 TTTGAGACTCAAGACAAAGATTTGTGCTTGGCAATTAAACAGCTTCCAAAGAGTGTG 1857  
Db 4975 TTTGAGACTCAAGACAAAGATTTGTGCTTGGCAATTAAACAGCTTCCAAAGAGTGTG 5034  
Qy 1858 CTTCCGATTAAACAGAACCTTGAATTGATGATCTTCTCTGCTTCCAGACTGTCCG 1917  
Db 5035 CTTCCGATTAAACAGAACCTTGAATTGATGATCTTCTCTGCTTCCAGACTGTCCG 5094  
Qy 1918 TATTTTGGGAAATTCGGGTGAGTGTCAAAAGGATCTTCCAAAGAGTGTGCTGAG 1977  
Db 5095 TATTTTGGGAAATTCGGGTGAGTGTCAAAAGGATCTTCCAAAGAGTGTGCTGAG 5154  
Qy 1978 GCATGCTCTGCTGCTCTCTATGATGCGGGATTAAGACCTTCAATTGAGGAGAGTGGAA 2037  
Db 5155 GCATGCTCTGCTGCTCTCTATGATGCGGGATTAAGACCTTCAATTGAGGAGAGTGGAA 5214  
Qy 2038 GATTTCTGCTCATGCTTTGGCAACCAACCACTGCGGAGCTGAGCTTGGGACAGAC 2097  
Db 5215 GATTTCTGCTCATGCTTTGGCAACCAACCACTGCGGAGCTGAGCTTGGGACAGAC 5274  
Qy 2098 ATCTGACAGAGCGGGCCATGAAGACCTGTGTGCAAGCTGAGGATCCCACTGCAAG 2157  
Db 5275 ATCTGACAGAGCGGGCCATGAAGACCTGTGTGCAAGCTGAGGATCCCACTGCAAG 5334  
Qy 2158 ATACAGACCTTGATTTTGAAGATCAGACATTAACCTCTGCTGACAGACTCTGAGAG 2217  
Db 5335 ATACAGACCTTGATTTTGAAGATCAGACATTAACCTCTGCTGACAGACTCTGAGAG 5394  
Qy 2218 ATGCTCATGCGCAACCGTAACCTTAAGATCCCTCAACTTGGAGGACCACTGAAGAA 2277  
Db 5395 ATGCTCATGCGCAACCGTAACCTTAAGATCCCTCAACTTGGAGGACCACTGAAGAA 5454  
Qy 2278 GAGAGTGTGAAGATGCGGTGGAAGCTTAAACCAACCAATTTTGTGAGAGCTTTG 2337  
Db 5455 GAGAGTGTGAAGATGCGGTGGAAGCTTAAACCAACCAATTTTGTGAGAGCTTTG 5514  
Qy 2338 AGGCTGAGTTGCTGAGATTGAACCAATGCTGTTAAGCTGAAGATCTCCCAATCTTACG 2397  
Db 5515 AGGCTGAGTTGCTGAGATTGAACCAATGCTGTTAAGCTGAAGATCTCCCAATCTTACG 5574  
Qy 2398 ACCCTCCCAAGCTTGAATCTTGAAGCTTGAAGGAAACAAGTGAAGACCAAGGAGTA 2457  
Db 5575 ACCCTCCCAAGCTTGAATCTTGAAGCTTGAAGGAAACAAGTGAAGACCAAGGAGTA 5634  
Qy 2458 ATGCTCTCATGATGATGCTTGAAGATCTTCCAGTGGCCCTTGAAGACTGATACTGAG 2517  
Db 5635 ATGCTCTCATGATGATGCTTGAAGATCTTCCAGTGGCCCTTGAAGACTGATACTGAG 5694  
Qy 2518 GACTGTGGCATCAACAGCGGTTTGCAGAGTGTGAGCTGAGCTGCTGTCAGCAACCGG 2577  
Db 5695 GACTGTGGCATCAACAGCGGTTTGCAGAGTGTGAGCTGAGCTGCTGTCAGCAACCGG 5754

QY 2578 AGCTTGACACACTGTGCTTATCCAAACAGCCCTGGGAAAGAGTGAATCTACTG 2637  
Db 5755 AGCTTGACACACTGTGCTTATCCAAACAGCCCTGGGAAAGAGTGAATCTACTG 5814  
QY 2638 TGTGCATCATGAGGCTTCCCACTGATGCTGCAGAGGCTGATGCTGAATCATGCTGAC 2697  
Db 5815 TGTGCATCATGAGGCTTCCCACTGATGCTGCAGAGGCTGATGCTGAATCATGCTGAC 5874  
QY 2698 CTGACAGAGGCTGAGTGTGTTTCTTGCATGCTGCTTAAGGCTGAATCATGCTGAC 2757  
Db 5875 CTGACAGAGGCTGAGTGTGTTTCTTGCATGCTGCTTAAGGCTGAATCATGCTGAC 5934  
QY 2758 CACCTGAGCCTTAGCATGAACCTCTGGAAGACATGAGCTGAAGCTTCTGTGCGAGGTC 2817  
Db 5935 CACCTGAGCCTTAGCATGAACCTCTGGAAGACATGAGCTGAAGCTTCTGTGCGAGGTC 5994  
QY 2818 ATGAGAGAACCATCTTGTATCTCCAGAGACCTGAGATTGTAAAGTGTATCTCACCGCC 2877  
Db 5995 ATGAGAGAACCATCTTGTATCTCCAGAGACCTGAGATTGTAAAGTGTATCTCACCGCC 6054  
QY 2878 GCGTCTGTGAGAGTCTGTCTGTGTATCTTCGAGAGACACACCTGAAGAGCTTGAT 2937  
Db 6055 GCGTCTGTGAGAGTCTGTCTGTGTATCTTCGAGAGACACACCTGAAGAGCTTGAT 6114  
QY 2938 CTCACGACCAATGCGCTGGGTGACGCTGGGCTGTGCGCTGTGCGAGGACCTGAAGCAA 2997  
Db 6115 CTCACGACCAATGCGCTGGGTGACGCTGGGCTGTGCGCTGTGCGAGGACCTGAAGCAA 6174  
QY 2998 AAGAACAGTGTCTGACGAGACTCGGCTTGAAGGCTATGACTGATCTGTGATTTGCTGT 3057  
Db 6175 AAGAACAGTGTCTGACGAGACTCGGCTTGAAGGCTATGACTGATCTGTGATTTGCTGT 6234  
QY 3058 GAGGACCTCTGCTGCTGCTTCTGCAACGCGGACTGACAGCTGAACCTGGTGGAG 3117  
Db 6235 GAGGACCTCTGCTGCTGCTTCTGCAACGCGGACTGACAGCTGAACCTGGTGGAG 6294  
QY 3118 AATACTTCAGTCCCAAGAGATGATGAGCTGTGCTGCTTGTGCTGCCACGCT 3177  
Db 6295 AATACTTCAGTCCCAAGAGATGATGAGCTGTGCTGCTTGTGCTGCCACGCT 6354  
QY 3178 AACTTACAGATTAATTTGG 3194  
Db 6355 AACTTACAGATTAATTTGG 6371

RESULT 10  
CQ731113 2753 bp DNA linear PAT 03-FEB-2004  
LOCUS CQ731113  
DEFINITION Sequence 17047 from Patent WO02068579.  
ACCESSION CQ731113  
VERSION CQ731113.1 GI:42306827  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kites, such as nucleic acid arrays, comprising a majority of  
humanecons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 17047 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
Source  
1. 2753  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 63.3%; Score 2553.8; DB 6; Length 2753;  
Best Local Similarity 94.1%; Pred. No. 0;

Matches 2736; Conservative 0; Mismatches 2; Indels 171; Gaps 1;  
QY 404 GACATGAGGTGACACATGGGACTTACAAGAGTCAAGTATGACCAAAATTCGCTGAGGAG 463  
Db 1 GACATGAGGTGACACATGGGACTTACAAGAGTCAAGTATGACCAAAATTCGCTGAGGAG 60  
QY 464 AGATGTACGTCGTAGTTTGAATAACAATGCTGCTGACATGCGGAAATGCAAACTTTGG 523  
Db 61 AGATGTACGTCGTAGTTTGAATAACAATGCTGCTGACATGCGGAAATGCAAACTTTGG 120  
QY 524 CTGCTGCTTTGATTCAGACCGGTGGGCTTCCGCGCTTCGACGCTGTTCTGCACGGA 583  
Db 121 CTGCTGCTTTGATTCAGACCGGTGGGCTTCCGCGCTTCGACGCTGTTCTGCACGGA 180  
QY 584 AGTCAGGAATTTGGAAATTCGCTTACGACGAAGATCGATGCTGCTGGCGGCAAGGTG 643  
Db 181 AGTCAGGAATTTGGAAATTCGCTTACGACGAAGATCGATGCTGCTGGCGGCAAGGTG 240  
QY 644 GACTCTACAGAGGATGTTCTCTACGCTCTTCTCTCTCCCGTTAGAGAGATGACGCGGA 703  
Db 241 GACTCTACAGAGGATGTTCTCTACGCTCTTCTCTCTCCCGTTAGAGAGATGACGCGGA 300  
QY 704 AGAAGAGAGCAGTGTCAAGAGTTCACTTCAGAGAGTGGCCAGACTCCAGGCTCCGG 763  
Db 301 AGAAGAGAGCAGTGTCAAGAGTTCACTTCAGAGAGTGGCCAGACTCCAGGCTCCGG 360  
QY 764 TGACGAGATGATGCTCCGACAGAAAGGCTGTTGATCATGATGACGCTTTCGATGAC 823  
Db 361 TGACGAGATGATGCTCCGACAGAAAGGCTGTTGATCATGATGACGCTTTCGATGAC 420  
QY 824 TGGGCTCTGCTCTCAACATGACCAAAAGCTCTGCAAAAGCTGAGGCTGAGAGAGCCTC 883  
Db 421 TGGGCTCTGCTCTCAACATGACCAAAAGCTCTGCAAAAGCTGAGGCTGAGAGAGCCTC 480  
QY 884 CGTTCAACCTTATACGAGTGTGCTGAGAGAGGCTGCTGCTGAGAGCTTCTGATGAC 943  
Db 481 CGTTCAACCTTATACGAGTGTGCTGAGAGAGGCTGCTGCTGAGAGCTTCTGATGAC 540  
QY 944 TCACCGTCAAGAGAGTGGGACAGAGAGCTCAAGTCAAGAGTGTGCTGCTGCTGCTGAC 1003  
Db 541 TCACCGTCAAGAGAGTGGGACAGAGAGCTCAAGTCAAGAGTGTGCTGCTGCTGCTGAC 600  
QY 1004 TGTAGTTAGAGGATCTCCGCGGACCAAAAGATCACTTGTCTTGAAGCGGAGATTG 1063  
Db 601 TGTAGTTAGAGGATCTCCGCGGACCAAAAGATCACTTGTCTTGAAGCGGAGATTG 660  
QY 1064 GTGACATCAGAACAACAGAGGCTTGGCTGATCATGAAACACCTGAGCTGCTCGAAC 1123  
Db 661 GTGACATCAGAACAACAGAGGCTTGGCTGATCATGAAACACCTGAGCTGCTCGAAC 720  
QY 1124 AGTGCAGGTGCGCGGTGGGCTCTGATCTGAGGTGGGCTGAGCTGAGAGAGCGTGG 1183  
Db 721 AGTGCAGGTGCGCGGTGGGCTCTGATCTGAGGTGGGCTGAGCTGAGAGAGCGTGG 780  
QY 1184 TGGGGAGAGGCTGCGCCCTTCAACCAAAAGCTCAAGGCTGACAGCGCTTTTGTGT 1243  
Db 781 TGGGGAGAGGCTGCGCCCTTCAACCAAAAGCTCAAGGCTGACAGCGCTTTTGTGT 840  
QY 1244 TTCAATCAGCTACCCCTGAGAGGCTGCTGCGGCTGCTCTCAATCTGAGAGAAAGATTG 1303  
Db 841 TTCAATCAGCTACCCCTGAGAGGCTGCTGCGGCTGCTCTCAATCTGAGAGAAAGATTG 900  
QY 1304 TCCGAGAGCGTCTTGGCGATATGCTGAGAGGAGTGGAAATGAGAAAGTCAAGTTTG 1363  
Db 901 TCCGAGAGCGTCTTGGCGATATGCTGAGAGGAGTGGAAATGAGAAAGTCAAGTTTG 960  
QY 1364 ATGTGACGACCTCATGATGTTCAAGAGCTCGGAGGCTGAGAGCTCGGCTCTGTTTCA 1423  
Db 961 ATGTGACGACCTCATGATGTTCAAGAGCTCGGAGGCTGAGAGCTCGGCTCTGTTTCA 1020  
QY 1424 TGAACATCTCTTCCAGAGACGCACTGTGAGAGTACTACACTTCTTCCACTCAGTGC 1483  
Db 1021 TGAACATCTCTTCCAGAGACGCACTGTGAGAGTACTACACTTCTTCCACTCAGTGC 1080



REFERENCE 3 (bases 1 to 157141)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (15-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
COMMENT On Jul 15, 2000 this sequence version replaced gi:7690109.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.sbgc.stanford.edu  
Quality: Phrap Quality >=40 99.9% of Sequence;  
Estimated Total Number of Errors is 0.1.

## FEATURES

source  
1..157141  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTC-490M1.0"

## ORIGIN

Query Match 39.6%; Score 1598.8; DB 9; Length 157141;

Best Local Similarity 99.9%; Pred.No. 0;  
Matches 1600; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 400 CAAGGACATGAGGTGACATGAGGACATCAAGAGTCACTGATGACCAAAATTCGCTGAG 459
DB 42003 CAGAGACATGAGGTGACATGAGGACATCAAGAGTCACTGATGACCAAAATTCGCTGAG 41944
QY 460 GAGAGAGATGATCGTCTGATGATTTGAAAAACCTGCTGCTGACTGCGGAAATGCAACG 519
DB 41943 GAGAGAGATGATCGTCTGATGATTTGAAAAACCTGCTGCTGACTGCGGAAATGCAACG 41884
QY 520 TTGGCTGCTGCTTTTGTGATTCAGACCGGTGGGCTTCGCGCTCGACGATGCTTGTGAC 579
DB 41883 TTGGCTGCTGCTTTTGTGATTCAGACCGGTGGGCTTCGCGCTCGACGATGCTTGTGAC 41824
QY 580 GAAAGATGAGAAATGAGAAATCGCTTACGACGAAAGATGATGCTGCTGGGCGCA 639
DB 41823 GAAAGATGAGAAATGAGAAATCGCTTACGACGAAAGATGATGCTGCTGGGCGCA 41764
QY 640 GGTGACCTTACCGAGAAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 699
DB 41763 GGTGACCTTACCGAGAAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 41704
QY 700 CGGAAAGAGAGAGAGAGTCAAGAGTCACTTCCAGAGAGTGGCGACCTCCAGGCT 759
DB 41703 CGGAAAGAGAGAGAGTCAAGAGTCACTTCCAGAGAGTGGCGACCTCCAGGCT 41644
QY 760 CCGGTGACGAGATCATGCTCCGACAGAAAGGCTGTTGTTCAATCATTTGACGTTTCGAT 819
DB 41643 CCGGTGACGAGATCATGCTCCGACAGAAAGGCTGTTGTTCAATCATTTGACGTTTCGAT 41584
QY 820 GACCTGGGCTTGTCTTCAAGATGACAAAGCTCTGCAAAAGACTGGGCTGAGAGCAG 879
DB 41583 GACCTGGGCTTGTCTTCAAGATGACAAAGCTCTGCAAAAGACTGGGCTGAGAGCAG 41524
QY 880 CCTCGGTTACCTCATAGAGTGTGCTGAGGAAAGTCCGCTCCGAGTCCCTTCTG 939
DB 41523 CCTCGGTTACCTCATAGAGTGTGCTGAGGAAAGTCCGCTCCGAGTCCCTTCTG 41464
QY 940 ATCGTCACTGACAGAGTGGGCAAGAGAGTCAAGTCAAGAGTGTGCTTCTCCGCT 999
DB 41463 ATCGTCACTGACAGAGTGGGCAAGAGAGTCAAGTCAAGAGTGTGCTTCTCCGCT 41404
QY 1000 TACCTGTTAGTTAAGAAATCTCCGAGGAAACAAAGATTCACCTGCTCTTGAAGCGGG 1059
DB 41403 TACCTGTTAGTTAAGAAATCTCCGAGGAAACAAAGATTCACCTGCTCTTGAAGCGGG 41344
QY 1060 ATTGTTGATGATCAAGAAACAAAGGCTTGGTGGATCAATGAACAAACCTGAGTGTGTC 1119
DB 41343 ATTGTTGATGATCAAGAAACAAAGGCTTGGTGGATCAATGAACAAACCTGAGTGTGTC 41284
```

```
QY 1120 GACCAGTGCAGAGTGGCCGCGTGGGCTCTCTCATCTGCGTGGCCCTGACGCTGACAGAC 1179
DB 41283 GACCAGTGCAGAGTGGCCGCGTGGGCTCTCTCATCTGCGTGGCCCTGACGCTGACAGAC 41224
QY 1180 GTGGTGGGAGAGCGTGGCCCTTCAACCAAGCTCACAGGCTGACGCGCTTTT 1239
DB 41223 GTGGTGGGAGAGCGTGGCCCTTCAACCAAGCTCACAGGCTGACGCGCTTTT 41164
QY 1240 GTGTTTATGACGCTCACCCCTCGAGGCGTGGTCCGCGGCTGCTCAATCTGGAGAAAGA 1299
DB 41163 GTGTTTATGACGCTCACCCCTCGAGGCGTGGTCCGCGGCTGCTCAATCTGGAGAAAGA 41104
QY 1300 GTTGTCTGAAGCGCTTCTGCGGATGAGCTGTGAGAGGAGTGGAAATGAGAACTGAG 1359
DB 41103 GTTGTCTGAAGCGCTTCTGCGGATGAGCTGTGAGAGGAGTGGAAATGAGAACTGAG 41044
QY 1360 TTTGATGATGACACCTCATGATTTCAAGAGACTCGGGAGTGTAGCTCCGCTCTGTTT 1419
DB 41043 TTTGATGATGACACCTCATGATTTCAAGAGACTCGGGAGTGTAGCTCCGCTCTGTTT 40984
QY 1420 CACATGAACATCTTCTCCGAGACAGGCACTGTGAGGAGTACTACACCTTCTCACCTC 1479
DB 40983 CACATGAACATCTTCTCCGAGACAGGCACTGTGAGGAGTACTACACCTTCTCACCTC 40924
QY 1480 AGTCTCAGAGCTTCTGCGCGCTTGTACTACGTTTGAAGGCGCTGGAATGAGCCA 1539
DB 40923 AGTCTCAGAGCTTCTGCGCGCTTGTACTACGTTTGAAGGCGCTGGAATGAGCCA 40864
QY 1540 GCTCTTGGCTCTGTGACGTTGAGAAACAAAGAGTTCATGAGCTTTAAACAGGCGGC 1599
DB 40863 GCTCTTGGCTCTGTGACGTTGAGAAACAAAGAGTTCATGAGCTTTAAACAGGCGGC 40804
QY 1600 TTGCATATCCACCTGCGTTTGGATGAGAGGTTTCTGTTTGGCTCGTGAAGCAAGCTA 1659
DB 40803 TTGCATATCCACCTGCGTTTGGATGAGAGGTTTCTGTTTGGCTCGTGAAGCAAGCTA 40744
QY 1660 AGAGGCGCATGAGAGTCTGCTGGGCTGTCCGCTTCCCTCGGGGCTGAAGCAGAAGCTT 1719
DB 40743 AGAGGCGCATGAGAGTCTGCTGGGCTGTCCGCTTCCCTCGGGGCTGAAGCAGAAGCTT 40684
QY 1720 CTGACATGGGCTCTCTGTTGGGTCAAGAGCTTAATGCAACACCCAGAGAGACCTTG 1779
DB 40683 CTGACATGGGCTCTCTGTTGGGTCAAGAGCTTAATGCAACACCCAGAGAGACCTTG 40624
QY 1780 GAGGCTTCCACCTGCTTTTTCAGAGCTCAAGACAAAGAGTTGTTGCTGGCAATTAAC 1839
DB 40623 GAGGCTTCCACCTGCTTTTTCAGAGCTCAAGACAAAGAGTTGTTGCTGGCAATTAAC 40564
QY 1840 AGCTTCCAGAAAGTGTGGCTTCCGATTAAACCAAGACTGGAATTGATAGCATCTTCTTC 1899
DB 40563 AGCTTCCAGAAAGTGTGGCTTCCGATTAAACCAAGACTGGAATTGATAGCATCTTCTTC 40504
QY 1900 TGCTTCCAGACCTGTCCGATATTGGGAAATTTGGGTGATGTCAAAGGATCTTCCCA 1959
DB 40503 TGCTTCCAGACCTGTCCGATATTGGGAAATTTGGGTGATGTCAAAGGATCTTCCCA 40444
QY 1960 AGAGATGAGTCCGCTGAGGAGCATGCTGTGGTCCCTGATGG 2001
DB 40443 AGAGATGAGTCCGCTGAGGAGCATGCTGTGGTCCCTGATGG 40402
```

## RESULT 12

AC024580 193609 bp DNA linear PRI 21-DEC-2001  
DEFINITION Homo sapiens chromosome 19 clone CTD-2621117, complete sequence.

AC024580  
AC024580.6 GI:17975240  
KEYWORDS  
SOURCE  
ORGANISM

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 193609)



AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 193609)  
AUTHORS DOE Joint Genome Institute.  
JOURNAL Direct Submission  
SUBMITTED (29-FEB-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 193609)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
JOURNAL Direct Submission  
SUBMITTED (13-NOV-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
4 (bases 1 to 193609)  
REFERENCE DOE Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS Direct Submission  
JOURNAL Submitted (21-DEC-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Dec 21, 2001 this sequence version replaced gi:16905144.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www-shgc.stanford.edu  
Quality: Phrap Quality >=40.99.4% of Sequence;  
Estimated Total Number of Errors is 0.9.  
FEATURES  
source  
1..193609  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTD-2621117"  
ORIGIN  
Query Match 39.6%; Score 1598.8; DB 9; Length 193609;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1600; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
400 CAGGACATGAGGTGACACATGGAAGCTACAGTATGACCAATTCGCTGAG 459  
22255 CAGGACATGAGGTGACACATGGAAGCTACAGTATGACCAATTCGCTGAG 22314  
460 GAGGAGATGATGCTGCTGATTTTGAACACCTGCTGACCTGCGGAAATGCAAG 519  
22315 GAGGAGATGATGCTGCTGATTTTGAACACCTGCTGACCTGCGGAAATGCAAG 22374  
520 TTGGCTGCTGCTTTGATTCAGACCGGTGGGCTTCGGCTCGACCGGTGCTGAC 579  
22375 TTGGCTGCTGCTTTGATTCAGACCGGTGGGCTTCGGCTCGACCGGTGCTGAC 22434  
580 GGAAGATCAGAAATGGGAATCGGCTTAGCCAGAAAGATCGTGTGCTGGGCGCA 639  
22435 GGAAGATCAGAAATGGGAATCGGCTTAGCCAGAAAGATCGTGTGCTGGGCGCA 22494  
640 GGTGACCTTACAGAGGAATGTTCTCTACGCTCTTCTCCCGTTAGAGATGCG 699  
22495 GGTGACCTTACAGAGGAATGTTCTCTACGCTCTTCTCCCGTTAGAGATGCG 22554  
700 CGGAGAGAGAGAGAGAGTGCACAGAGTTCATCTCCAGGAGTGGCAGACTCCAGGCT 759  
22555 CGGAGAGAGAGAGAGAGTGCACAGAGTTCATCTCCAGGAGTGGCAGACTCCAGGCT 22614  
760 CCGGTGACGAGATCATGTCCGACCCAGAAAGCTGTGTTGTTATCATTTGACGGTTCAT 819  
22615 CCGGTGACGAGATCATGTCCGACCCAGAAAGCTGTGTTGTTATCATTTGACGGTTCAT 22674  
820 GACCTGGGCTCTGTCTCTCAACAATGACCAAAAGCTTCGAAAGAGTGGGCTGAGAGCAG 879  
22675 GACCTGGGCTCTGTCTCTCAACAATGACCAAAAGCTTCGAAAGAGTGGGCTGAGAGCAG 22734  
880 CTTCCGTTCAACCTCATAGCAGAGTCTGCTGAGAAAGTCTCTGCTCCCTGAGTCTTCTG 939  
22735 CTTCCGTTCAACCTCATAGCAGAGTCTGCTGAGAAAGTCTCTGCTCCCTGAGTCTTCTG 22794

940 ATTCGACCGTACAGAGAGTGGGCAACAGAAAGCTCAAGTACAGAGTGTGCTCCCGT 999  
22795 ATTCGACCGTACAGAGAGTGGGCAACAGAAAGCTCAAGTACAGAGTGTGCTCCCGT 22854  
1000 TACCTGTAGTTAGAGATCTCCGGGAAACAAAGATTCACCTTCTCTTGAAGCGCGG 1059  
22855 TACCTGTAGTTAGAGATCTCCGGGAAACAAAGATTCACCTTCTCTTGAAGCGCGG 22914  
1060 ATTGATGACATCAGAAACACAGAGTGGGCTGCTGATCATCAACACCGTGAAGTCTGCT 1119  
22915 ATTGATGACATCAGAAACACAGAGTGGGCTGCTGATCATCAACACCGTGAAGTCTGCT 22974  
1120 GACGATGCGAGATGCTCCCGGCTGGGCTCTCATCTGCTGGGCTGCTGAGTGAAGAC 1179  
22975 GACGATGCGAGATGCTCCCGGCTGGGCTCTCATCTGCTGGGCTGCTGAGTGAAGAC 23034  
1180 GTGATGGGAGAGAGTGGGCTCCCTTCAACCAAGCTCACAGGCTGACCGCTTTT 1239  
23035 GTGATGGGAGAGAGTGGGCTCCCTTCAACCAAGCTCACAGGCTGACCGCTTTT 23094  
1240 GTTTCATCAGCTCACCCCTTCCAGGCGTGTCCGCGCTGTCTCATCTGAGGAAAGA 1299  
23095 GTTTCATCAGCTCACCCCTTCCAGGCGTGTCCGCGCTGTCTCATCTGAGGAAAGA 23154  
1300 GTTTCCTGAGAGCGCTTCTGCGCTATGCTGTGAGAGGAGTGTGAATGAGAGTCACTG 1359  
23155 GTTTCCTGAGAGCGCTTCTGCGCTATGCTGTGAGAGGAGTGTGAATGAGAGTCACTG 23214  
1360 TTTGATGATGAGAGCTCATGATGTTCAAGAGTCCGAGAGTCTGAGCTCGTCTGTTT 1419  
23215 TTTGATGATGAGAGCTCATGATGTTCAAGAGTCCGAGAGTCTGAGCTCGTCTGTTT 23274  
1420 CACATGAACATCTCTTCCAGACAGCACTGTGAGAGTACTACACCTTCTTCCAGCTC 1479  
23275 CACATGAACATCTCTTCCAGACAGCACTGTGAGAGTACTACACCTTCTTCCAGCTC 23334  
1480 AGTCTCAGAGACTTCTGTCGCGCTGTGATCAAGTGTGAGAGGCTGTGAATGAGAGCA 1539  
23335 AGTCTCAGAGACTTCTGTCGCGCTGTGATCAAGTGTGAGAGGCTGTGAATGAGAGCA 23394  
1540 GCTCTGCGCTCTGATGTTGAGAGCAAAAGAGTCCATGAGACTTAAACAGGCAAGC 1599  
23395 GCTCTGCGCTCTGATGTTGAGAGCAAAAGAGTCCATGAGACTTAAACAGGCAAGC 23454  
1600 TTCCATATCACTGCTTGTGATGAGAGGCTTCTGTTGAGCTGTGAGGCAAGACGTA 1659  
23455 TTCCATATCACTGCTTGTGATGAGAGGCTTCTGTTGAGCTGTGAGGCAAGACGTA 23514  
1660 AGAGGCGCATGAGAGTCTGCTGGGCTGTCCCGTTCCTGGGGGTGAGCAAGCTT 1719  
23515 AGAGGCGCATGAGAGTCTGCTGGGCTGTCCCGTTCCTGGGGGTGAGCAAGCTT 23574  
1720 CTGCACTGGGCTCTCTGTGGGTGAGAGGCTTATGCAACCAACCCAGAGACACCTG 1779  
23575 CTGCACTGGGCTCTCTGTGGGTGAGAGGCTTATGCAACCAACCCAGAGACACCTG 23634  
1780 GACGCTTCACTGCTTGTGAGAGTCAAGCAAAAGTGTGTTGCTTGGCATTAAC 1839  
23635 GACGCTTCACTGCTTGTGAGAGTCAAGCAAAAGTGTGTTGCTTGGCATTAAC 23694  
1840 AGCTTCAAGAGATGCTGCTTCCGATTAAACCAAGACTGAGACTGTATGATCTTCTTC 1899  
23695 AGCTTCAAGAGATGCTGCTTCCGATTAAACCAAGACTGAGACTGTATGATCTTCTTC 23754  
1900 TGCCCTCAGCACTGTGCTGATTTGGGAAATTTGGGGGTGATGATGAGAGTCTTCCA 1959  
23755 TGCCCTCAGCACTGTGCTGATTTGGGAAATTTGGGGGTGATGATGAGAGTCTTCCA 23814  
1960 AGAGATGAGTCCGCTGAGAGATGCTGCTGTGATCTCTATGAG 2001  
23815 AGAGATGAGTCCGCTGAGAGATGCTGCTGTGATCTCTATGAG 23856



RESULT 13  
LOCUS AY721594 3545 bp mRNA linear MAM 15-SEP-2004  
DEFINITION Bos taurus maternal antigen that embryo require mRNA, complete cds.  
ACCESSION AY721594  
VERSION AY721594.1 GI:51980124  
KEYWORDS  
SOURCE .  
ORGANISM Bos taurus (cow)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.  
REFERENCE 1 (bases 1 to 3545)  
Penneker, S., Uzbekova, S., Perreau, C., Papillier, P., Merrillod, P. and Dalbès-Tran, R.  
Spatio-Temporal Expression of the Germ Cell Marker Genes MATER, ZARI, GDP9, BMP15, and VASA in Adult Bovine Tissues, Oocytes, and Preimplantation Embryos  
Biol. Reprod. (2004) In press  
2 (bases 1 to 3545)  
Penneker, S., Uzbekova, S. and Dalbès-Tran, R.  
Direct Submission  
Submitted (12-AUG-2004) Physiologie de la Reproduction et des Comportements, Institut National de la Recherche Agronomique, Nouzilly 37380, France  
FEATURES  
SOURCE location/Qualifiers  
1..3545  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
1..3297  
/note="MATER; NALP5"  
/codon\_start=1  
/product="maternal antigen that embryo require"  
/protein\_id="AAU20764.1"  
/db\_xref="GI:51980125"  
/translation="MKRAKIAPIALSNYGLQMFQIQKEEPQTKALLKEHASESAQS  
FPLVQVADADASLSLHEHCRASLAKTSDIPEKMSALSMSMADDEKKYLIAE  
ISDPSAPKTDGQSGMEKVPGRPDPOSRIRIHWMTPTSPTRDTPRPFEPSECP  
DAHALSGAFNPDPSGFRPLTVLHGPVGKSSIARLLILFMAQGDILKYLPSYFL  
LRARDLQSRRTSFALILSKEMPDAPVEVNLQPEELLIVQGLELELTFPDQDS  
SLADWAEKQAPVLAHSLKLVLLPECALILTVQDAGLQRLQALRSPRLVWGSL  
VENRMQLLIGGKCRKRTCAHAGADHEDVLKQGVVVALVEALELGEPEGKL  
PVGHTLLGLYATFVFORLAPDAGMRALSGEBREGALGLRIADGYWNAKPYDGD  
DLGVHGLQGPRLSALQASILLPOGHCGRGAFSLILSQEPFALFVLRVEDDGG  
YPLPFQSTKSLTELRHIDLNVQLVQMKRFLGLVSKVEYRALLETLLGCPVPAKQQL  
LHWI CLVGOHPAAASPDLBAFYCLFEAODEFVRLALNGFQEWMLQINPMDLTVS  
SFCLRRCOHLRRKRLVDVGTDPDEFAEMSGAPQGLKXTIDHEVEDLSYLSTHPLN  
RQLDLSGSLVLSKEAMKTLGVKLRQPAKCIQMLIFKGAIVTGLRHLMTLLINNTIR  
LDLNGCRLEEDVOTACALHHPQCALSLDLDRGLPASCRLSIQVLAATSGSLKSL  
SLTGKLVADQGVKSLCDALKVTPCTLOGLIISGCGLTAAQODLALILEMOGLTHLS  
LSGDLGSKGMSLDCRAVKLSSCGLOKALANASLDVAGCCFLVAFALMGNNHTLHLS  
SMNPLEDPEGNMLCEVMEPSCPLKDLVNCRLTASCKSLSNVITRSPRLSDLA  
ANALGDEGILALCEGLKQKNTLTRLGLRACGLTSEGCALANALITCSRHLSLIMRN  
DLGPGMTLLCSAFMHPTSNQITIGLMREQYPARVRLLEQVORLKPHVIVSDAWYTE  
EEEDGPCWRI"  
1..291  
/note="Region: pyrin/DAPIN domain"  
538..1047  
/note="Region: NACHT domain"  
2035..2118  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
2203..2286  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
2290..2373  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
2374..2457  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"

misc\_feature 2461..2544  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
misc\_feature 2545..2628  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
misc\_feature 2632..2715  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
misc\_feature 2716..2799  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
misc\_feature 2803..2886  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
misc\_feature 2887..2970  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
misc\_feature 2971..3054  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
misc\_feature 3055..3138  
/note="Region: ribonuclease inhibitor subtype leucine  
repeat"  
3' UTR  
polyA\_signal  
polyA\_site  
ORIGIN  
Query Match 38.8%; Score 1566; DB 4; Length 3545;  
Best Local Similarity 68.4%; Pred. No. 0;  
Matches 2271; Conservative 0; Mismatches 1005; Indels 42; Gaps 6;  
1 ATGAAGAGCAATATGCTCACTTTCCAGCTACGAGGCGTCGATGCTCTATAG 60  
1 ATGAGAGAGGCCAAATAGCCCCCTTATCCATTACGACGTGCTGTTGAGCAG 60  
QY 61 CTAGACAAGAGAAATTTACAGACATTCAGAGAAATTAAGAAATTTCCAGAAATCG 120  
DB 61 CTGGGCAAGAGAGATTCACAGACATTTAAGCCCTCTGMAAGAGACATGCTCAGAAATCA 120  
QY 121 ACCACATGCTCTATTTCCACAGTTTGAATCGAATCGAACAGTGTCTGGCACTC 180  
DB 121 GCACAGTGCTCTTCCGCTGTGTCAAGTGAACAGGCTGATGCGAGTCCCTGCCCTCC 180  
QY 181 CTCTTGATGATTTATGAGCAATCGTGGCCCTGGGCTAGTCCATTAGATTTTGAA 240  
DB 181 CTCTGATGAGATTTATGAGCAATCGTGGCCCTGGGCTAGTCCATTAGATTTTGAA 240  
QY 241 AACATGAACCTGCGAACCCCTCTCGAAGAGGACGGGATGACATGAAAAAATTTCAAA 300  
DB 241 AAGATGAGCTGTGGGCACTGTCTGAGATGACAGGACGAGATGAAAAATATCTACTG 300  
QY 301 GCTATGAACAGAAAGGTGCCACAGCAGCAGACAGACAGAAACAAAAATTTCAAAAGCT 360  
DB 301 GCTAAATATACAGAAATTTCTGACCAACAAAGAC----- 335  
QY 361 ATGAACAGAAAGGTGCCACAGCAGCAGACAGACAGAAACAAAGACATGAGGTGACACA 420  
DB 361 --TACCAAGGTCCAAGATGAAGAAAGTGCAGGCCCCAGAGAGACCCACAGAGACTCG 393  
QY 421 TGGACTTACAGAGTCAAGTATGACCAAAATTCCTGAGAGAGAGATGTACGTGTAAGT 480  
DB 421 TTTGAAGAAATTTGCTGAGATGTCCGAGCGCGATGCGCTATCGGGGGCTTTAAACCA 513  
QY 481 TTTGAAGAACTGCTGCTGATGAGCGCGAAATGCAAAAGTTGAGTGTGCTTTGATTCA 540  
DB 481 TTTGAAGAAATTTGCTGAGATGTCCGAGCGCGATGCGCTATCGGGGGCTTTAAACCA 513  
QY 541 GACCGGT---GGGCTTCGGGCTTCGACAGGTGTTCTGACCGGAAAGTCAAGAAATGGG 597  
DB 541 GACCCCTCGGGGGGCTTCGGGCTTCGACAGGTGTTCTGACCGGACCCCGGGGGGTGGG 573

QY 598 AATCGGCTCTAGCCAGAAAGATCGTGTGTGCGGCGAAAGTGAATCTACAGGGA 657  
 DB 574 AGTCTCGCTGGCCCGGAGGCTCTACTGTTTCTGGGCGGAGGCACTCTACAAAGGCG 633  
 QY 658 ATGTTCTCTAGCTCTTCTTCTCTCCCGTTAGAGATGCAAGCGAAGAGAGACAGT 717  
 DB 634 TTGTTCTCTAGCTCTTCTCTCTCCCGCGCAGAGACCTCCAGGGGCTCAGGAGAGACAG 693  
 QY 718 GTCAAGATTTATCTCCAGGAGGTGGCCAGACTCCAGAGCTCCGTTAGAGAGATCATG 777  
 DB 694 TTTCGCGAGCTCATCTCCAGAGAGTGGCGGAGCGCCCGGTCGCGTGAAGAGAGTCTG 753  
 QY 778 TCCCGACAGAAAGCGTGTGTGTCTATCATTTAGCGTTCGATGACCTG---GCTCTGTC 834  
 DB 754 TCCAGACTCGAAAGGCTCTTGAATGTGTGATGATGACTGGAAGAGCTGAGACTCACCTTC 813  
 QY 835 CTCAACATGACACAAAGCTCTGCAAAAGCTGCGGCTGAGAGAGCAAGCTCCGTTCACTTC 894  
 DB 814 AGAGACAGAGATCCAGCTCTCGCTGCTGAGCTGGGCGGAGAGGAGCGCCCGCCGTCGTC 873  
 QY 895 ATACGAGTCTGTGAGAGAGTCTGCTCTCTGAGTCTCTGATGCTCACCTGACAG 954  
 DB 874 GCGCAGAGCTGTGAGAGAGTCTGCTCTCCGAGTGGCGCTCTCTCTCACTCTGTCAG 933  
 QY 955 GACGTGGGACAGAGAGCTCAAGTCAAGTCAAGTCTCTCCCGTTACCTGTTAGTTAGA 1014  
 DB 934 GACGCGGAGTTGACAGAGGCTCCAGGCTCTCGCTCCGTTCTCCGTTACTATGAGTCTG 993  
 QY 1015 GGAATCTCCGCGGAAACAAAGATCCACTTGTCTCTTGAAGCGCGGAGTTGATGAGATCAG 1074  
 DB 994 GGCCTCTCAGTGGAAAAGAGATGACAGTTGCTCTCGGCGGCGGAGAGACTGCGCTGCT 1053  
 QY 1075 AAGACACAGGCTGTGCTGCTGATCATGAACACGTTGAGTCTGACCACTGCGCAGTGG 1134  
 DB 1054 AAGACATGCGCTGTGACCGCGGCGCGCAGACACAGAGAGTGTCTTACAGTCTCAGTGG 1113  
 QY 1135 CCGCGCGTGGGCTCTCATCTGCGGCGCTGCGAGCTGACAGAGCTGTTGGGAGAGAC 1194  
 DB 1114 CCGTGTGTGTGCGCTGTGCTGTGAGGCTCTGAACTGACAGGAGAGCGGAGAGAGAGC 1173  
 QY 1195 GTGCGCCCTTCAACCAACGCTCAAGGCTGACAGCGCTTGTGTTCATCAGCTC 1254  
 DB 1174 CTTCCGCTCCCGGCGCACACCTCAAGGCTTGTAGCCACCTCTGCTTCCAGCGGCTG 1233  
 QY 1255 ACCCTCGAGGCGTGTGCGGCGCTGTCTCAATCTGAGAGAAAGATGTTCTTGAAGCGC 1314  
 DB 1234 GCTCCCAAAGATGACAGGCTGCGCGCGCTGAGCGGAGAGAAAGCGGTCCTCGAAGGAGC 1293  
 QY 1315 TTCTGCGCTATAGCTGTGAGAGAGATGTGTGAATAGGAATCAGTGTGTTGATGTGACGAC 1374  
 DB 1294 TTGTGCGGCTGTGACAGCGGACGCGGTGTGAACGCGAATTCGTGTGTTACGCGACGAC 1353  
 QY 1375 CTATGAGTTCAAGACTCGGAGAGTCTAGCTCCGCTCTGTTTCACTGAACATCTT 1434  
 DB 1354 CTGGGCGTCTCAAGGCTGACAGGCGCCGAGCTCTCGGCTCTGACAGAGCGAGCATCTT 1413  
 QY 1435 CTCCAGACAGCACTGTGAGAGATCTACACTTCTTTCACCTCAGTCTCAGAGACTTC 1494  
 DB 1414 CTCCCTGACGCGCACTGCGGAGAGGCGCACGCAATTCCTCCACCTGAGCTCCAGGAGTTC 1473  
 QY 1495 TGTGCGCTTGTATCACTGTGTTAGAGGCGCTGGAATGAGCGAGCTCTGCGCTCTG 1554  
 DB 1474 TTGTGCGCTTGTATCACTGTGCGAGGCGTGTGAGGAGAGCGGAGGAGCTTACCGGCTG 1533  
 QY 1555 TACGTTGAGAGACAAAGAGTCAATGAGCTTAAACAGGAGGCTTCACTATTCACCTG 1614  
 DB 1534 TTCCCGCAGAGACAAAGAGTCTGACAGAGCTCAAGCATATTTAGCTTCAAGCTTCAAGCTG 1593  
 QY 1615 CTTTGATGAGAGCTTCTTGTGTGCGCTGTGAGGAGAGAGATGAGAGGCACTGAG 1674  
 DB 1594 GTCCAGATGAGAGATCTTATTTGGCTTGTGAGAGAGAGATGAGGCGGCTTGGAA 1653  
 QY 1675 GTCCGTGTGGGCTGTCCGTTCCCTGCGGAGGTGAAGCAGAACTTGTGACTGCGCTCT 1734

DB 1654 ACCCTCTAGGCTGTGCTCCGTTGCGCCCGGTGGCTAAGCAGCACTTGTGACTGATCTGCG 1713  
 QY 1735 CTGTTGGGTCAAGAGCTTAATGCTCAACCCAGAGAGACCTCTGAGCGCTTCCACTGT 1794  
 DB 1714 CTGTGGGTCAAGAGCTTGTGCTCCGCTCCGAGACTTGTCTGAGGCTCTTACTGTC 1773  
 QY 1795 CTTTTCAGACTCAAGAGCAAAAGTGTGTGCTTGGCACTTAAACAGCTTCCAGAGTGG 1854  
 DB 1774 CTTTTCAGAGCCAGAGACAGAGTTCGTTGCTGTGGCTTGAACGAGCTTCCAGAGTGG 1833  
 QY 1855 TGGCTTCCGATTAACAGAACTTGATGATCTTCTGCTGCTCCAGAGACTGT 1914  
 DB 1834 TGGCTGCAATTAACCGGCAATGACATTAACGAGTCTCTTCTGTCTCCGCGCTTGC 1893  
 QY 1915 CCGTATTTGCGGAAAATTGCGGTGATGTCAAAAGGATCTTCCAGAGATGATCGCT 1974  
 DB 1894 CAGCATTTTACGAAAGTTGCACTGATGTCAAGAGGA---CCCAAGAGATGAATTTGCT 1950  
 QY 1975 GAGGATGTCTGTGTGCTCTTATGATGTGGGATTAAGACCTCATTTAGAGAGAGTGG 2034  
 DB 1951 GAGGATGTCTGGGCTCTCCAGAGGCTGAAGATCAAGACCC---TTGATGAGACCTGG 2007  
 QY 2035 GAAAGTTTCTGCTCATGTGCTTGGCACCACCCAGACTGCGGAGCTGGAAGCTGGGAGC 2094  
 DB 2008 GAAAGCTTGTCTCTGCTGTGCTGACGACGACCCAACTTACAGAGCTGACCTGAGTGGC 2067  
 QY 2095 AGCATCTGACAGAGCGGCGCATGAAGACCTGTGTGCGCAAGCTGAGGATCCACTGC 2154  
 DB 2068 AGCTCTTGAGCAAGAGGCGCATGAAGACCTGTGTGCTGCAAGCTGCGGAGCAGCTGT 2127  
 QY 2155 AAGATACAGACCTGATGTTTGAATGCAAGATTAACCTGTGTGTGAGACCTCTGG 2214  
 DB 2128 AAAATACAGAACTGATTTTAAAGGCGCGCGGTTACCCCTGTCTCGTACCTCTGG 2187  
 QY 2215 AGAATGTATGAGGCAACCGTAACTTAAGATCCCTCACTTGGAGAGGCCACCACTGAAG 2274  
 DB 2188 ATGACTGTGATTAATCAACCGAATCAACGCGCTGTGACGTGACGAGCTGCGCTGAGA 2247  
 QY 2275 GAAAGGATGTAAAGATGGCGGTGTGAAGCTTAAACACCCAAATGTTTGTGTGAGTCT 2334  
 DB 2248 GAGAGAGAGCTGTCAACCGCGGTGTGAGGCGCTCAAGGACCCACAGTGTGCGTGAAGTCT 2307  
 QY 2335 TTGAGGCTGATGTGTGTGATTAACCAATGCTGTAACTGTAAGATCTCCAAATCTT 2394  
 DB 2308 CTGAGGCTTGAATGCTGTGTGATTAACCCAGCTCTTGTGCGGAGATCTCCCAATGCTT 2367  
 QY 2395 AGCACTCCCGAGCTGAATCTGTGAGCTGTGAGAGAGAAACAAGTGAAGAGACAGAGGA 2454  
 DB 2368 GTTACGTCCGAGCTGAATCTGTGAGCTTACGAGCTTACGAGAAATGAAGTGTGAGAGG 2427  
 QY 2455 GTATGCTCTTCAAGATGAGCTTGTGAGAGTCTTCCAGTGTGCGCTGACAGAGCTGATCTG 2514  
 DB 2428 GTGAAGTCTCTGTGAGAGCTTGAAGTCAACCTTGCACCTTGCACAGAGTAACTCTG 2487  
 QY 2515 GAGAGCTGTGATCAAGAGCAAGGCTTGCAGAGTGTGAGCTGTGAGCTGTGAGCAAC 2574  
 DB 2488 GGGAGCTGTGAGCTCAAGCGGCACTGCGCAAGAGCTGTGCTGTCTCATCAGAGAC 2547  
 QY 2575 CGAGCTTGAACACCTGTGTCTATCAACCAAGCTGTGGGAGAGAGAGTGAATCTA 2634  
 DB 2548 CAGGCTTGAACACCTGTGTCTGTGCGGAGATGAACCTGGGAGAGAGAGTGAAGCTG 2607  
 QY 2635 CTGTGTGATCATGAGGCTTCCCACTGTATGTGACAGAGCTGTGATCATGCTG 2694  
 DB 2608 CTGTGTGCGGCTGTGAAGCTCTCAGCTGTGTGTGTGAGAGAGTGTGACATTAACCATG 2667  
 QY 2695 CACTGACAGAGCTGT 2754  
 DB 2668 AGCTGTGATGT 2727  
 QY 2755 ACGCACTGAGCTTGAAGTAACTGTGTGAAGCAATGCGTGAAGCTTGTGTGTGTGTGTGT 2814

Db 2728 ACGCACTGAGCTCAGCATGAAACCCCTGAGAGATCCCGAGATGAACTTCTGTGCGAG 2787  
Qy 2815 GTCAATGAGAGAACCAATCTTGTATCTCCAGAGACCTGAGTTGGTAAGTGCATCTCAC 2874  
Db 2788 GTCAATGAGAGAACCAATCTTGTATCTCCAGAGACCTGAGTTGGTAAGTGCATCTCAC 2847  
Qy 2875 GCCCGCTCTGAGAGAGTCTGTCTGTGTGATCTTGAAGAGACGACACCTGAGAGCTG 2934  
Db 2848 GCCTGCTGTGAGAGAGTCTGTCTGTGTGATCTTGAAGAGACGACACCTGAGAGCTG 2907  
Qy 2935 GATCTCAAGAGACAAATGCTTGGTGTGAAGGAGTGTGCTGCTGCTGTGAGAGAGCTG 2994  
Db 2908 GATCTGAGAGAGACAAATGCTTGGTGTGAAGGAGTGTGCTGCTGCTGTGAGAGAGCTG 2967  
Qy 2995 CAAAGAGACAGTGTCTGTGAGAGACTCGGAGTGTGAAGGAGTGTGAGTGTGATTTG 3054  
Db 2968 CAGAGAGACAAATGCTTGGTGTGAAGGAGTGTGCTGCTGCTGTGAGAGAGCTG 3024  
Qy 3055 TGTGAGAGACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3114  
Db 3025 TGCAG 3084  
Qy 3115 CAG 3174  
Db 3085 CGCAAG 3144  
Qy 3175 TCTAATCTCAAGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3234  
Db 3145 TCTAATCTCAAGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3204  
Qy 3235 CTGAG 3294  
Db 3205 CTGAG 3264  
Qy 3295 GATGAG 3312  
Db 3265 GAG 3282

RESULT 14  
AY329487 3405 bp mRNA linear ROD 18-DEC-2003  
LOCUS AY329487 Mus musculus strain CBA/J MATR protein isoform-E (Mater) mRNA,  
DEFINITION complete cds; alternatively spliced.  
ACCESSION AY329487  
VERSION AY329487.1 GI:37595472  
KEYWORDS  
SOURCE  
ORGANISM Mus musculus (house mouse)  
MUS musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS 1 (bases 1 to 3405)  
TITLE Identification and characterization of alternative splice variants  
for murine Mater  
UNPUBLISHED  
2 (bases 1 to 3405)  
AUTHORS Cheng, H., Huang, Z., Zhang, Y., Liu, H., Teuscher, C. and Ma, R. Z.  
JOURNAL Direct Submission  
TITLE Submitted (20-JUN-2003) Human and Animal Genetics Division,  
Institute of Genetics and Developmental Biology, Chinese Academy of  
Sciences, Building 917, Datun Road, Beijing 100101, P.R.China  
FEATURES  
SOURCE  
1..3405  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CBA/J"  
/db\_xref="taxon:10090"  
1..3405  
/gene="Mater"  
9..3296  
/gene="Mater"  
/note="alternatively spliced"

/codon\_start=1  
/product="Mater protein isoform-E"  
/protein\_id="AA094610.1"  
/db\_xref="GI:37595473"  
translation="MGPEKDSKAILKARLEGEKSEKMTSPENDSKISIQKQGE  
QEQTSITWPEKDSKAILKARLEGEKSEKMTSPENDSKISIQKQGE  
ERKMTSPENDSKISIQKQGEQEQTSITWPEKDSKAILKARLEGEKSEKMTSPENDSKISIQKQGE  
LSDAFPEKQTFPHITLILGRPEVGSALARSITVLMAQOKLEOKS FVIFSVREI  
KMTESKSLAOLIAKECPDSMDPVTKIMSOPRLLIFVLDGDDMSVLOKDDMTSRDW  
KDBPIYILNWSILKRALPDSPLIITTRTTEKLSKMSVSPILYILVEGLSARSQ  
LVLENISNEDRIYFHSILBNHOLFQOCAPSCSYVCEALDQKLGKRTIPCT  
LTGLYATLVPHQLTKRPSOSALSQEQITLVGICMAAEQVMTSRVFDVDDLKNS  
LKSESIILAFPMNILLQVHNSQCYFSLSLDDFPAALYYVLEGEENQKCFE  
NORSIMEVKRTDTRILKMRFLGLMNDIKLTLEVLFEYPIVPEYOKLQHWLSI  
AOVNGTSPMDTDAFYCLFESODEEYVGLKROEVLILINQMDLKVSXYCLKIC  
ONKATIRVDIROLISVNTLETCPVTVVOETOCKEPLIMEMWNGCSYLGLSRKEID  
LGDSILSORAKIICLELRNOSCHTQULTKESAVSVGLKMLKLSNQKLTNLG  
NTPMKDDMKLACALHAPKCSVETLLDSCETLIGYEMISTLLISTRLKLSLAK  
NRVGVKMSISLGNALSSMCLQKLIIDNCGELFASCHLIVSALFSNQNLTHLCLSN  
SLTEGVQOLCOFLANPECALORLIINHCVDAYGFPLARLANTKLTHLSLTPNP  
VGDGAMLLCEALKEPTCYLOEILVDCOLTONCCEPLACMITTTKOLKSLDLGNML  
GDKVITLCEGLKONNSLRRLGAKCKLTSNCEALSLAISCPHINSILNTLNKDS  
TSGMLKICSAFQCEVSNLGIITLWKQRYAVRQLEVEFEKHHVVIDGWDVASED  
DRNWKQ"

## ORIGIN

Query Match 31.9%; Score 1287.8; DB 10; Length 3405;  
Best Local Similarity 65.4%; Pred. No. 3.1e-307;  
Matches 1985; Conservative 0; Mismatches 1032; Indels 18; Gaps 6;

Qy 287 AAAAAATTTCAAGAGTATGAGAACAGAGAGTGCACAGCAGCAGAGACAGAGAACAGAACAG 346  
Db 256 AAAAAATTTCAAGAGTATGAGAACAGAGAGTGCACAGCAGCAGAGACAGAGAACAGAACAG 315  
Qy 347 AATTTCAAGAGTATGAGAACAGAGAGTGCACAGCAGCAGAGACAGAGAACAGAACAG 406  
Db 316 AAAAAATTTCAAGAGTATGAGAACAGAGAGTGCACAGCAGCAGAGACAGAGAACAGAACAG 375  
Qy 407 ATGAGAGTGAACAGAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAG 466  
Db 376 AGCAGAGAGAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGA 435  
Qy 467 ATGAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAG 526  
Db 436 ATGAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAG 495  
Qy 527 GTGCTTTGATTCAGACCGGTGGGGCTTCCGGCTCGACCGGTGGGTTCGACGAGAACGT 586  
Db 496 ATGCTTTTAAACCAATTCAGAGAACCTTCCAGCTTCACACCATTAATCTTACATGAGAAC 555  
Qy 587 CAGGATTTGGGAATTCGCTTACGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 646  
Db 556 CAGGATTTGGGAATTCGCTTACGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 615  
Qy 647 TCTACAGAGAGATGCTTCTTACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 706  
Db 616 TCTTCCAAAAATG---TCTTGTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 672  
Qy 707 AGGAGAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAG 766  
Db 673 AGGAGAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAG 732  
Qy 767 CGGAGATGATGCTCCGAGAGAGAGT 826  
Db 733 CAAAGATATGCTCCAGAGAGAGAGT 792  
Qy 827 GCTCTGTCTTC---AAGATGACACAAAGCTCTGCAAAAGCTGGGCTGAGAGAGAGCTTC 883  
Db 793 ACTCTGTCTTCACACAGATGATATGACATATCAAGAGCTGGAAGATGAACAGAGCTCA 852  
Qy 884 CGTTCACTCTATACAGAGT 943  
Db 853 TATCAATCTGATGATACAGCTCTGAGAGAGAGCTTCTTCACTTCTTCTTCTTCTTCTTCTTCA 912

QY	944	TCACCGTCAGAGACGTGGGCACAGAGAGCTCAACATCGAGAGTGTGTCTCCCGTTACC	1003
Db	913	TTACACCAAGAAACAAGGCTTAAAGAAAATCTAAAGTCAATGGTGTCTTCCCGCTCTATA	972
QY	1004	TGTTAGTTAGAGGAATCTCCGGGGAAACAAAGATCACTTTCCTCTTAGCGCGGAGATTG	1063
Db	973	TACTGTTGAAGGACTGTCTGTGATCAAGAGATCTCACTGTGTCTCGAAGAAATCTCCA	1032
QY	1064	GTGAGCATCAGAAAGACAAAGGTTGCGTGCGATCATGAACAACCGTAGCTCTCGACC	1123
Db	1033	ATGATCTGATAGAAATACAAAGCTTCCATTTCTGATAGAAAATCAACCAAGCTGTGTGACC	1092
QY	1124	AGTGCAGAGTGCACGCGCGGGCTCTCATCTGTGGGCCCTGACGTGACAGGCGTGG	1183
Db	1093	AATGCGAGGCCCTCTGTGTGTCTCCCTGTGTGTGAAGGCTCTCGAGCTTACAGAAAGAAC	1152
QY	1184	TGGGGAGAGCGTGCAGCCCTTTCACAAACGCTCACAGGCTGCACGCGCTTGTGTGT	1243
Db	1153	TGGGAAAGATGACACCTTACCCTCCAGACTCTCACGCGTTGTGATGCCAGTTGGTGT	1212
QY	1244	TTTCATCACTCAACCCCTCGAGGCGGTGTCCGCGCTGTCTCAATCTGAAGAAAGATTTG	1303
Db	1213	TTCAACACCTCACTTGAAGAAAGGCTTCTCCAGAGGCTCTCAGTACGAGAAACAGATTAA	1272
QY	1304	TCCTGAACGCTTCTGCGCCGTATAGCTGTGAAGGAGTGTGAATAGAAATCAAGTGTGG	1363
Db	1273	CTCTAGTGGTTTTGTGCATGATGCGAGCTGAAGAGATGTGACATAGAGTCCGTGTCT	1332
QY	1364	ATGTGACGACCTCATGCTTTCAAGGACTCGGGAGTCTGAGCTCCGTCTCTGTTCACA	1423
Db	1333	ATGAAGATGATCTGAAGAACTAATAGCTTAAAGGATCTGAATCTTGGCCCTCTTTCACA	1392
QY	1424	TGAACATCTTCTCCACAGACGCC--ACTGAGGAGATATCAACCTTCTCCACCTCA	1480
Db	1393	TGAACATCTTCTCCAGGTTGGCCACAAACAGTAGACATGTATATTTTCTTCCACCTCA	1452
QY	1481	GTCTTCAGGACTTCTGTGCGCGCTTGTATCTACGTTTAGAGGCGCTGGAATTCAGCCAG	1540
Db	1453	GCTTCAGGATTTCTTGTCTGTGCTTATATTTATTTTGAAGGGCTGGA--GGATGGA	1509
QY	1541	CTCTTGCCTCTGTACGTGTGAAGAAACAAAGAGTCCATGAGCTTAAACAGCAGCT	1600
Db	1510	ATCAGCAATTTTGTCTTCAATTGAAAACAAAGAGCATATGAGAGGTGAAGAACTGACG	1569
QY	1601	TCCATATTCACCTCGCTTTTGAATGAAGACGTTTTCTTGTGGGCTCGTGAGCGAAGCGTAA	1660
Db	1570	ACACTCGCTC---CTTGGGATGAAGCTTTCTTATTTTGGCTCTATGAACAAGATATCT	1626
QY	1661	GGAGGCACTGAGAGTCTGTGCTGGCTGTCCGTTCCTTGGGGGTGAAGCAAGCTTC	1720
Db	1637	TGAAGACTCTGGAGGTTCTGTTTGAATATCCCGTATTCCAACTGTTAGCAGAGCTCC	1686
QY	1721	TGCATGTGGTCTCTGTGTGGTGACGAGCCTAATGCCAACCCAGAGACACCTCG	1780
Db	1687	AACACTGGGATCTCTGTAGTACGACAGCTCAATGAGCAACAGCCCAATGACACCTCG	1746
QY	1781	ACGCTTCACACTGTCTTTTCGAGACTCAAGACAAAGATTTGTTCCGCTTGCATTAAACA	1840
Db	1747	ATGCTTTCTATGTCTTATTTTGAAGTCTCAGATGAAGAGTTTGTGGCGGGCTCTCAAC	1806
QY	1841	GCTTCCAGAAATGTGGCTTCCGATTAACAGAACTTGGACTTGTATAGACTTTCTTCT	1900
Db	1807	GCTTCCAGAAAGTGTGCTGTGATTAACAGAAAGATGACCTTGAAGTCTCTTCTTAACT	1866
QY	1901	GCTTCACAGACTGTCCGTATTTTGCGAAATTTCCGGTGGATGTCAAAAGGATCTTCCAA	1960
Db	1867	GTTCTAACACATGTAGAACTTTGAAGGCAATCCGGTGTGATATCAGAACTCTCTCTCG	1926
QY	1961	GAGATAGTCCGCTGAGGATGTCTGTGTGCTCT--CTATGATGCGGATTAAGACC	2017
Db	1927	TAGATAAATCTCTCAAGCTGTGCTCTGTGTATCTGTCTACAGAGACAAATGTAGGCCCT	1986

QY	2018	TCATTGAGGACACGTGGGAAATATTTCGCTCCATGCTTGGGACCCACCCACCTCGGCG	2077
Db	1987	TCCTCATGGAATGGGGGAACTTCGTCTGTGTGCTTGGCAGGCTCCGGAACCTTGAAAG	2046
QY	2078	AGCTGACCTCTGGGAGCAGCATCTCTGCACAGCGGGGCCATGAAGACCTGTGTGCACAG	2137
Db	2047	AGCTGGAATTGGGGCAGCAGCATCTTGAGTCAACGGGCCATTAAGATTACTGTGTCTCGAGC	2106
QY	2138	TGAGGCATCCCACTGCGAAGATACAGACCTGTATGTTTAAATGCAAGATTACCCCTG	2197
Db	2107	TGCGGAATCAGTCTCGCAGATACAGAAAGCTGAGGTTTAAAGATGACAGAGGTAGTGTG	2166
QY	2198	GTGTGCAGCACCTCTGAGGAATTCGTATGCGCAACCGTAACCTTAAGTCCCTCAACTTGG	2257
Db	2167	GCCTGAAACATCTCTGGMAAGCTCCCTTTTAAACAATCAAAACTTAAAGTAACTCAATCTAG	2228
QY	2258	GAGGACCCCACTGAAGGAAAGAGATGTAAAGATGGGGTGTGAAGCCTTAATACACCCCA	2317
Db	2227	GGAACTCTCCATGAAAGGATGATGACATGAATTAGCTTGCAAGCCGTGAACATCTCCA	2288
QY	2318	AATGTTTGTGGAGTCTTTTGAAGCTGATTTGCTGTGATTTGACCAATGCTCTTTACTTGA	2377
Db	2287	AGTGTCCGTGTGAGACTCTGAGTTTGTGATTCCTGTAGTTAACATATCATTTGGCTATGAGA	2346
QY	2378	AGATCTCCCAATCTCTTACGACCTTCCCGACCTGAAATCTGTAGCTGGCAGAAACA	2437
Db	2347	TGATCTCCACGCTCTTATTTTCAACCAACAGGCTAAAGTGTCTTACGCTCGCCAAAATA	2406
QY	2438	AGGTGACGACCAAGGAGATATGCTCTCAAGTATGCTTGAGAGTCTCCAGTGGCGCC	2497
Db	2407	GAGTGGAGTAAAGAATCATGATATTCCTTGGAAATGCTTGAGTACCTCAATGTGTCTAC	2466
QY	2498	TGCAAGAGCTGATCTGAGAGACTGTGGCATCACAGCCAAGGTTTGCAGAGTCTTGCGCT	2557
Db	2467	TGCAAAAGTTGATTACTGGACAACTGTGGCCCTCAACCTGCGACGCTGCACCTTGTGCT	2526
QY	2558	CAGCCCTCTGTAGCAACCCGAGCTTTGACACACTGTGCTATTCACAACAAGCCTTGGGA	2617
Db	2527	CAGCCCTTTTACGAAACCAAGAACTTGCACACCTGTGCTCTGCAAAACAAGCCTTGGGA	2586
QY	2618	ACGAAGGTGTAATCTTACTGTGTGATTCATGAAAGCTTCCCACTGTATGTGTGAGAGGC	2677
Db	2587	CTGAAGGATGTCACAAGCTGTGTGATTCCTGAGAAATCCAGAAATGTCTCTCAAGGCGC	2646
QY	2678	TGATGCTGAATCAGTGCCACCTGAGACAAGCTGTGCTGTGTGTTCTTGCACCTTGCGCTTA	2737
Db	2647	TGATATCTGAATCACTGCACAATGTGATGATGTCTTATGCTTCTTGCGCAATGAGACTTG	2706
QY	2738	TGGGTAACTCATGCTGACGCACTGACCTTACATGAACCTGTGTGAAGACAAATGAGCG	2797
Db	2707	CAAAACAACAAGACTGACCCACTGAGCGTGAACATGAACCCGTTAGGGGATGTGTGCA	2766
QY	2798	TGAAGCTTCTGTGGAAGTCAATGAGAAACATTTGTCAATCTCAAGACCTGGAATTTGG	2857
Db	2767	TGAAGCTACTGTGTGAAGCTTTAAAGAACCTTATTTACTTTCAAGAACTGGAACCTAG	2826
QY	2858	TAAAGTGCATCTACACCGCGCGTGTGTGAGACTGTCTCTGTGTATCTCGAGAGCA	2917
Db	2827	TGACTGCGCACTACACAGAACTGTGTGCGAGAACCTGTGCTGTATGATACAAACCA	2886
QY	2918	GACACCTGAAGAGCTGATCTCAACGACAAATGCCCTGTGGTGAAGCGTGGGGTTGCTGCG	2977
Db	2887	AGCAATTAATAAGTTTGGATCTTGATTAACAACGCCCTGGGGTGAACAAAGAGTCATTAACC	2946
QY	2978	TGTGTGAGGGACTGAAGCAAAAGAACAGTGTCTTGACAGACCTGGGTTGAAGCATGTG	3037
Db	2947	TGTGTGAGGGACTGAAGCAAAATTAACAGCTTCCCTGAAGAGACTTGGGTGTGGGGCATGTA	3006
QY	3038	GACTGACCTTCGATTTGTGTGAGGAGCACTTTCCTTGGGCCCTTCTGCAACCGGATCTGA	3097
Db	3007	AGTTGACTTCAATTTGCTGTGAGGCAATTTGTATTTGGCCATCTCTTTCGAACCCCTCACTGA	3066
QY	3098	CCAGTCTAAACCTGTGTGACAAATTAATTCAGTCCCAAGAAATGATGAAGCTGTGTTCGG	3157

Db 3067 ACAGCTAACCTGGTGAAGATGACTTCAGTACATCGGGAGATGTTGAAGCTGTCTCTG 3126  
Qy 3158 CCTTGCTGTCACGCTCTAATACATTAATGGCTGTGGAATGGAGTACCTG 3217  
Db 3127 GCTTCAATGACCTGTCTTAACCTGGAGATTAATGCTGTGGAAGAGAGATTAATG 3186  
Qy 3218 TGCATAATAGAGAGCTGTGAGAGAGTCACTCAAGCCCGAGTCTGTAATGAGC 3277  
Db 3187 CCCAGTGAAGAGAGAGCTGTGAGAGAGTGAATTTGTCTCAAGCCCACTGTGATTAATG 3246  
Qy 3278 GTAGTGGCATCTTTTATGATGAATGACCGAC 3312  
Db 3247 GTAGTGGTATGATGATGAATGACCGAAC 3281  
RESULT 15  
LOCUS AY329484 3432 bp mRNA linear ROD 18-DEC-2003  
DEFINITION Mus musculus strain SWR/J MATER protein isoform-E (Mater) mRNA,  
complete cds; alternatively spliced.  
ACCESSION AY329484  
VERSION AY329484.1 GI:37595466  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 3432)  
AUTHORS Cheng, H., Zhang, Y., Huang, Z. and Ma, R.Z.  
TITLE Identification and characterization of alternative splice variants  
for murine Mater  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 3432)  
AUTHORS Cheng, H., Huang, Z., Zhang, Y., Liu, H., Teuscher, C. and Ma, R.Z.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUN-2003) Human and Animal Genetics Division,  
Institute of Genetics and Developmental Biology, Chinese Academy of  
Sciences, Building 917, Datun Road, Beijing 100101, P.R.China  
FEATURES  
source location/Qualifiers  
1..3432  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="SWR/J"  
/db\_xref="taxon:10090"  
1..3432  
/gene="Mater"  
9..3296  
/gene="Mater"  
/note="alternatively spliced"  
/codon\_start=1  
/product="MATER protein isoform-E"  
/protein\_id="AA094607.1"  
/db\_xref="GI:37595467"  
/translation="MGPEKSKALKARGLEEKSEKSTPSKNTSPENDSKIOKOGPE  
OEOSTESTGPEPEKSKALKARGLEEKSEKSTPSKNTSPENDSKIOKOGPE  
ERKMTSPENDSKIOKOGPEOSTSPENDSKIOKOGPEOSTSPENDSKIOKOGPE  
LSDAPKYOXKQPTTILHGRPGVSKALARSIVLGAOKLQKMSFVLPFSREI  
KMTKSLAOLIAKSPDSMDLVTKMSOPERLFTVIGLDMDSVLQHDMTLSRDM  
KDEPITYLMTSLKALKPOSFLITTRNTGLEKLMVSPLYIIVELGASASRQ  
LVLEINSESDRIQVFSILBNHOLFDCQAPSVCLSCVCEALQLOKLGKSCCTPCOT  
LTGLVATLVFHLTKRPSOALSOEBITVLGCMMAAEVMTSRSVYDDLKNTS  
LKSEIILAFHNLILLOVGHNSBOCYVSHSIDPFALVYVLEEGSEMMOFCFIE  
NORIMEYKRTDTRLLGMKRFPLRGLMKDILKTEVLPYRPVPTPGQKQHWYSLI  
AQOVNGISPMOTLDAFYCLFSDSEFVGAALKRQEWLLINQKMDLVSSYCLKHC  
QNLKAIKVIDLSDVNTLECPVAVTQEQKPLMWMWNGFCSVGLSKRLKELD  
LGSILISORAMKILCLBLNOSCRIOKLPKSAEVSGLKLMKLFENOMLYLNLG  
NTPMKDDMKLACEALKPKKSVERTLDSCELFIIIGEMISTLLISTRLKSLIAK  
NRVGYKMSISLGNLSSMCLQKLLNCGLTTPASCHLVYVLEEGSEMMOFCFIE  
SLGTRGVQOLCQPLANPECALORLILNCHNVDAAYGLARLANTLTLTSLTNP  
VGDGAKMLCEALKSPCTCYLOBLELVQDLTONCEEDLACMTTTHKLSLDLGNAL  
GDKGVITLCEGLKQSSSLRLIGLGCALTSNCCBLALAI SCNPHLSLNUVNDPS  
TSGMLKCSAFQCPSVNLGIIIGLWKQEVYARVRQLBEVEFKPHVIVIGDMYASDED

ORIGIN DNMWKN"  
Query Match 31.9%; Score 1286.2; DB 10; Length 3432;  
Best Local Similarity 65.4%; Pred. No. 7.7e-307;  
Matches 1984; Conservative 0; Mismatches 1033; Indels 18; Gaps 6;  
Qy 287 AAAAAATTTCAAGCTATGGAACAGAGAGGTGCGACAGCAGACAGACAGAAACAG 346  
Db 256 AAAATGTCAGTAGAGCAATCTGAAAGACAGTGAATCAGAAAGAGTGAACAGCGCTAG 315  
Qy 347 AAATTTCAAGCTATGGAACAGAGAGGTGCCACAGCAGACAGACAGAAACAGAGAC 406  
Db 316 AAAAGAAAATGATCTTCTCCAAAAACAGACAGTAATATCATTCAGAAAGACCAAGACAG 375  
Qy 407 ATGAGAGTGAACACATGGGACTACAAAGTCAAGTCAAGTGAACCAATTCGTGAGAGAGAG 466  
Db 376 AGCAGAGACAGACATCAGATATATGAGGTGACTTACAAAGACTTACAGAGCCATGATTTG 435  
Qy 467 ATGTACGTCGTATGTTGAAACACCTGCTGACCTGCGCGGAATGCAACGTGTGCTG 526  
Db 436 CTAAAGTTCAGACAAAGTGTGATCTACACTATGACAGCCAGAGATGAATTTATTTGTG 495  
Qy 527 GTGCTTTGATTCAGACCGGTGGGCTTCGCGCTTCGACGCGTGTGTCAGCGAAAGT 586  
Db 496 ATGCTTTTAAACATACACAGAAACCTTCAGCTCACACCATTTATCTCATGAGAAAGAC 555  
Qy 587 CAGGAATGGGAATTCGCTCTACAGAGAGATGTGTGTGTGCGGCGGAGAGTGGAC 646  
Db 556 CAGAGATTGGAAAGTCAAGCTTTGCGCAGAAATATTTGTTGGCTGGGCAAGGATTAAC 615  
Qy 647 TCTACAGAGGAATGTTCTCTACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 706  
Db 616 TCTTCCAAAAAATG---TCTTTGTCACTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 672  
Qy 707 AGGAGAGAGTGTACAGAGATTCATCTCCAGAGAGTGGCCAGACTCCAGGCTCCGGTGA 766  
Db 673 AGAAGAGAGATTTGGACAGCTGATTTGTAAGAGGTGCCAGACTCTCGGATCTAGTGA 732  
Qy 767 CGGAGATCATGTCGCCAGACAAAGAGCTGTGTTCAATCATTTGAAGAGTGTGATGACCTGG 826  
Db 733 CAAGATCATATGTCACACAGAAAGACTCTGTTGTCAATGATGAGCTTGGATGATATAGG 792  
Qy 827 GCTCTGTCTC---AACAATGACCAAAAGCTCTGCAAAAGCTGGAGCTGAGAGAGCGCTC 883  
Db 793 ACTCTGTCTCACAATGATATATGACATATCAAGAGACTGAAGAGATGAACGCCCA 852  
Qy 884 CGTTACCTCATACGACAGTGTGCTGAGAAAGTCTGCTCTCTGAGTCTTCTGATCG 943  
Db 853 TATACATCTGATGTACAGCTCTCTGAGAAAGCTCTCTACCTCAAGTCTTTCATCA 912  
Qy 944 TCACCGTCAGAGAGCTGGGGCACAGAAAGCTCAAGTCAAGAGTGTGTCTCCGTTACC 1003  
Db 913 TTACCAACAGAAACAGAGCTTAAAGAAACCTCAAGTCAATGTTGTGCCCCCTCTATTA 972  
Qy 1004 TGTAGTTAGAGGAATCTCCGGGAAACAAAGATCCATTGCTCTTGGAGCGCGGAAATTG 1063  
Db 973 TACTGTGTTGAAGACTGTCTGCAATCAAGAGAAATCTCAAGTGTCTCTGAGAAACATCTCCA 1032  
Qy 1064 GTGAGCATCAGAAACACAGAGGTTGCGGTGATCATGAACAACCGTGAAGTCTCGAC 1123  
Db 1033 ATGAGTGTGAATGAATCAAGTCTTCAATCTGATTAAGAAATCAACAGAGCTGTTTGAAC 1092  
Qy 1124 AGTGCAGAGTCCCGCGGTGGCTCTCTCATCTGCGGTGGCCCTGCAAGTGTGACAGACGTGG 1183  
Db 1093 AATGCAGAGCCCTCTGTGTGCTCTCTGCTGTGAGGCTCTCAAGTCAAGAAAGAAAC 1152  
Qy 1184 TGGGAGAGAGGCGGCGCCCTTCAACAAAGCTCAAGGCTGCAAGCGCTCTTGTGTGT 1243  
Db 1153 TGGGAAAGATGACACCTTACCTGCGAGACTCTCAACGAGTTTGTATGACAGCTGGTGTGT 1212  
Qy 1244 TTCACTACAGCTACCCCTCGAGGCGTGTGTCTCGGCGCTGTCTCAATCTGAGAGAAAGTTG 1303

Db 1213 TTCAACAGCTCACTTGAAGAGGCTTCCAGAGCGCTCTCAGTCAGAGAAACAGATT 1272  
 Qy 1304 TCCGTAAGCGCTTCCGCTATGCTGTGGAGGAGTGTGAAATAGAAAGTCAAGTTTG 1363  
 Db 1273 CTTAGTGGGTTTGTGATGATGAGGAGCTGAAAGAGTGTGACATGAGTGGGTTCT 1332  
 Qy 1364 ATGTGAGCACTCACTGTTTCAAGAGCTGGGAGTGTGAGCTCGGTCTGTTTCA 1423  
 Db 1333 ATGATGATGATGAAAGAACTATAGCTTAAAGAGTGTGAGTCTTGCCCTTTCA 1392  
 Qy 1424 TGAATCTCTTCTCCAGACAGCC--ACTGTGAGAGTACTACACTTCTTCACTCA 1480  
 Db 1393 TGAATCTCTTCTCCAGGTTGGCCACAACAGTGAAGAGTGTATGTTTCTCCACTCA 1452  
 Qy 1481 GTCTCCAGAGCTTGTGTGGGCTTGTACTAGTGTAGAGGCTTGGAAATCCAGCCG 1540  
 Db 1453 GCTTCAGAGATTTCTTGTGCTTATATATATATATATATATATATATATATATAT 1509  
 Qy 1541 CTCTCTGCTCTGACGTTGAGAGACAAAGAGTCCATGAGCTTAAACAGCAGGCT 1600  
 Db 1510 ATCAGACTTTTGTCTTCAATGAAACCAAGAGCATCTGAGGTGAAAGAGACTGAC 1569  
 Qy 1601 TCCATATCCACTCGCTTGTGATGAAGCGTTTCTTGTGAGCTCGTGAACGAAGTAA 1660  
 Db 1570 ACACTCGCTC--CTTGGGATGAAGCGTTTCTTATTTGGCTCATGAAACAGATATCT 1626  
 Qy 1661 GAGGCGCACTGAGGCTCTGCTGGGCTGTCCGTTCCCTGGGGGTGAAGCGAAGCTTC 1720  
 Db 1627 TGAATCTCTGAGGTTCTGTTTGAATATCCGTAATCCATGCTGTGAGCAAGAGCTCC 1686  
 Qy 1721 TGCACTGGGCTCTGTTGGTTCAGAGCGCTAATGCGACCAACCCAGAGACACCTGG 1780  
 Db 1687 TAACATCTGGTCTCTGATAGCTCAGCAGGTCAATGAGCACAGCCCAATGAGACCTGG 1746  
 Qy 1781 AGCGCTTCACTGTCTTTTCAAGACTCAAGACAAAGATTGTGCTTGGCTTGAACA 1840  
 Db 1747 ATGCTTGTATGTCTATTTGAGTCTCAGGATGAAGAGTTGTGTGGGGGCTCTCAAC 1806  
 Qy 1841 GCTTCAAGAGTGTGGCTTCCGATTACAGAACCTGAGCTTGAATGATCTTCTTCT 1900  
 Db 1807 GCTTCAAGAGTGTGGCTGCTGATTAACAGAAATGAGCTTGAAGGCTCTTCTTCT 1866  
 Qy 1901 GCTTCAGAGCTGTGCTATTTGGGAAATTCGGGTGAGTGTCAAGGAGATCTTCCAA 1960  
 Db 1867 GTCTAAGACTGTCAAGCTTGAAGGCAATCGGGTGAATACAGAGACTCTCTCTGG 1926  
 Qy 1961 GAGATGAGTCCGCTGAGGATGCTGTGTGCTCT--CTATGATGCGGAGTAAAGCC 2017  
 Db 1927 TAGATATATCTCGAGGCTGTGCTGTGTGTTACTGTCCAGAGACAAATGTAAGCCC 1986  
 Qy 2018 TCATTTAGAGAGAGTGTGGAAATTTCTGTCTCATGCTTGGACCCACCACTGCGCG 2077  
 Db 1987 TCCATGATGAGTGTGGGAAATCTTGTCTGTGTGCTTGGCAGCTCCGGAATTTGAAG 2046  
 Qy 2078 AGCTGAGCTGTGGGAGAGCACTCTGACAGAGAGGGGCAATGAAGCCCTGTGTGCAAGC 2137  
 Db 2047 AGCTGAGCTGTGGGAGAGCACTCTGAGTCAACGCGGCATGAAGATCTGTGAGC 2106  
 Qy 2138 TGAGGCACTCCACTGCAAGATACAGACCTGATGTTTGAATCAACAGATTACCCCTG 2197  
 Db 2107 TCCGGAATCAAGTCTGTGAGAAATACAGAGCTGACGTTTGAAGTGTGAGAGTGTCTG 2166  
 Qy 2198 GTGTGAGAGCTGTGAGAAATGTCATGAGCCCAACCTGAACCTTAAGATCCCTCACTGG 2257  
 Db 2167 GCTTGAATCATCTGTGAAGCTCTTGTATGCAATCAAACTTAAGTACTCAATCTAG 2226  
 Qy 2258 GAGGACCACTGTAAGAGAGAGATGAAGATGCGCGTGAATGAGCTTAAACACCCAA 2317  
 Db 2227 GGAACATCTCCATGAAGAGATGAATGAATGAGTGAAGCTGCGCAAGCGCTGAACCTCA 2286  
 Qy 2318 AATGTTGTGAGAGTCTTGAAGCTGAGATTGCTGTGATGAGCCCATGCTGTACTGA 2377  
 Db 2287 AGTGTCCGTGAGACTGTGAGTGTGATCTGTGAGTAAACATCATGTTGTATGAGA 2346

Qy 2378 AGATCTCCAAATCTTACAGCTTCCCGAGCTGAATCTGTGAGCTGGAGAAACA 2437  
 Db 2347 TGAATCTCAAGCTTCTTATTTTCAACACAGCTTAAAGTGTCTGAGCTGGCAAAATA 2406  
 Qy 2438 AGGTGACAGACGAGAGATATGCTCTCAGTATGCTGAGAGTCTCCAGTGGCC 2497  
 Db 2407 GAGTGGAGTAAAGATGATATCTCTGGGAATGCCCTTGTGATGCTCAATGTGTAC 2466  
 Qy 2498 TGCAGAGCTGATATCTGAGAGCTGTGATCAACGCAAGGAGTTGCCAGATCTGGCT 2557  
 Db 2467 TGCAAAGTGTATCTGACAACTGTGCTCTCACACCTGCGAGCTGCCACTTGTGCT 2526  
 Qy 2558 CAGCCCTGTGACAGACGAGCTTGAACACTGTGCTTATCAACAAACAGCTGGGA 2617  
 Db 2527 CAGCCCTTTCAGCAACAGAACTTGACACCTGTGCTGTCAACAAACAGCTGGGA 2586  
 Qy 2618 ACGAAGTGAATCTACTGTGTGATCAATGAGCTTCCCACTGTAGTGTGACAGAGC 2677  
 Db 2587 CTGAAGAGTGAACAGCTGTGTGATCTGTGAGGAATCCAGAAATGTCTCCAGCGCC 2646  
 Qy 2678 TGAATCTGAATCAGTGCACCTGACACAGCTGTGTGTGTTTCTTGAACCTTGGCTTA 2737  
 Db 2647 TGAATCTGAATCAGTGCACCTGACACAGCTGTGTGTGTTTCTTGAACCTTGGCTTA 2706  
 Qy 2738 TGGATATCTAGTGTGAGAGACCTGAGCTTATGATGAACCTGTGAGAACATGAGCG 2797  
 Db 2707 CAACAAACAAAGTGAACCACTGAGCTGACCATTAACCCGTGAGAGATGTCGA 2766  
 Qy 2798 TGAAGCTTGTGAGAGCTCATGAGAGAACATCTTGTATCTCCAGAGCTGAGATTGG 2857  
 Db 2767 TGAAGCTTGTGAGAGCTTAAAGAACCTTGTATCTTGAAGATCTGAACTGAGACTAG 2826  
 Qy 2858 TAAAGTGTATCTCACCGCGGTGTGTGAGAGTCTGTCTGTGTGTCTGAGAGACA 2917  
 Db 2827 TGAATCTGCACTCAACAGAACTGTGAGAGACTGAGAGCTGAGTGTATCAACAAACA 2886  
 Qy 2918 GACACTGAAGAGCTGTGATCTCACAGACATGCTGTGTGTGTGTGTGTGTGTGTGT 2977  
 Db 2887 AGCACTTAAAGATTTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2946  
 Qy 2978 TGTGAGAGAGCTGAAGAGAAAGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3037  
 Db 2947 TGTGAGAGAGCTGAAGAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3006  
 Qy 3038 GACTGACTTGTGATGT 3097  
 Db 3007 AGTTGACTTCAATGT 3066  
 Qy 3098 CCAATCTTAACCTGT 3157  
 Db 3067 ACGACTTAAACCTGT 3126  
 Qy 3158 CTTTGTGCTGT 3217  
 Db 3127 GCTTCAATGCTGT 3186  
 Qy 3218 TGCAAATGAAGAGCTGT 3277  
 Db 3187 CCCGAGTGAAGAGAGCTGT 3246  
 Qy 3278 GTAGTGTGATCTTGT 3312  
 Db 3247 GTAGTGTGATCTGATGATGAAGATGACCGAAGC 3281

Search completed: July 18, 2005, 22:13:41  
 Job time : 17042 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2005, 12:25:31 ; Search time 2009 Seconds  
(without alignments) 1189.575 Million cell updates/sec

Title: US-10-066-521-5  
Perfect score: 4035  
Sequence: 1 atggaagaggaagcaatgcgtc.....tcgaggggctgtgtctttaa 4035

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4035	100.0	4035	6	AAL44356 Human PYR
2	4035	100.0	4035	10	ADF94744 Human PYR
3	3166.2	78.5	3900	6	ABK48628 Human MAT
4	3166.2	78.5	3900	8	AAD49018 Human MAT
5	3139	77.8	3830	8	ADA45220 Human MAT
6	3133.6	77.7	3489	6	AAD41224 Human EMB
7	3133.6	77.7	3926	8	ADA45218 Human MAT
8	3057.8	75.8	5859	6	AAL47135 Human dom
9	3057.8	75.8	6939	6	AAL47131 Human dom
10	3057.8	75.8	6939	6	AAL47140 Pyrin dom
11	2900.6	71.9	3226	6	ABX97181 Human NOV
12	1840.6	45.6	2820	12	ADN62078 Human CDN
13	1278	31.7	3447	8	ABK48611 Mouse MAT
14	1278	31.7	3447	8	AAD49001 Mouse MAT
15	1152.8	28.6	1157	6	ABK48609 Human MAT
16	1152.8	28.6	1157	8	AAD48999 Human MAT
17	756.6	18.8	1075	6	ABK48610 Human MAT
18	756.6	18.8	1075	8	AAD49000 Human MAT
19	536	13.3	2808	12	ADP47735 Human inf
20	527.8	13.1	3039	12	ADP47733 Human inf

21	440	10.9	2775	12	ADO50889 Human NOV
22	438	10.9	2775	10	ADF44704 Human NOV
23	424	10.5	2985	6	ADK36408 Human PAA
24	424	10.5	2985	12	ADJ19330 Human PAN
25	422.4	10.5	2982	10	ADF94764 Human PYR
26	422.4	10.5	3190	10	ADA53511 Human cod
27	422.4	10.5	3368	6	AAL44366 Human PYR
28	422.4	10.5	3368	10	ADP94760 Human PYR
29	418.2	10.4	2575	6	AAL170683 Human nuc
30	418.2	10.4	2767	8	ABT16018 Human relA
31	419.2	10.4	3172	6	ABN99366 Human sec
32	419.2	10.4	3310	13	ACN40690 Tumour-as
33	418.8	10.4	2883	12	ADO41637 Novel hum
34	409	10.1	3186	6	AAL44363 Human PYR
35	409	10.1	3186	10	ADP94754 Human PYR
36	404.4	10.0	3108	6	ADK36416 Human PAA
37	404.4	10.0	3218	12	ADJ19338 Human PAN
38	404.4	10.0	3218	12	ADJ19382 Human PAN
39	404.4	10.0	3300	6	AAL47129 Human PAN
40	404.4	10.0	3731	12	ADP47723 Human Mon
41	397.2	9.8	2799	12	ADH18983 Human cel
42	394.2	9.8	3531	6	AAL47128 Pyrin dom
43	392.6	9.7	3189	5	AAL67185 Nucleotid
44	392.6	9.7	4944	10	ADP14464 Human end
45	381.4	9.5	3431	4	ADJ14322 Human nuc

## ALIGNMENTS

RESULT 1	
AAL44356	
ID AAL44356 standard; CDNA; 4035 BP.	
XX	
AC AAL44356;	
XX	
DT 31-OCT-2002 (first entry)	
XX	
DE Human PYRIN-5 cDNA sequence.	
XX	
KW Human; gene; ss; gene therapy; PYRIN; stress-related response;	
KW Apoptotic response; inflammatory response; inflammatory disorder;	
KW Immune system disorder; Crohn's disease; multiple sclerosis; cancer;	
KW Leukemia; autoimmune disorder; arthritis; neurological disease;	
KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;	
KW Tissue typing; forensic biology; predictive medicine; pharmacogenomics;	
KW transcription profiling; PYRIN-5.	
XX	
OS Homo sapiens.	
XX	
FN Key	Location/Qualifiers
FT CDS	1..4035
FT	/*tag= a
FT	/product= "Human PYRIN-5"
XX	
WO200261049-A2.	
XX	
PD 08-AUG-2002.	
XX	
PF 31-JAN-2002; 2002WO-US002967.	
XX	
PR 31-JAN-2001; 2001US-0265231P.	
XX	
PR 10-SEP-2001; 2001US-0318645P.	
XX	
PA (MIL-) MILLENNIUM PHARM INC.	
XX	
PI (AMHP ) WYETH.	
XX	
FI Bertin J, Wang W, Blatcher M,	
XX	
DR WPI; 2002-627477/67.	
XX	
DR P-PSDB; AAO15585.	
XX	
PT New PYRIN polypeptides and nucleic acids useful for modulating and	





Db 1741 GGTGACGACCTTAATGCCACCAACCCAGAGACACCTTGAGCGCTTCCACTGTCTTTTC 1800  
Qy 1801 GAGACTCAAGAACAAAGAGTTTGTGCTTGCAATTAACAGCTTCCAGAAATGTGGCTT 1860  
Db 1801 GAGACTCAAGAACAAAGAGTTTGTGCTTGCAATTAACAGCTTCCAGAAATGTGGCTT 1860  
Qy 1861 CCGATTAAACAGAACCTTGACTTGTATAGCATCTTCTGCTCAGACAGTGTCCGTAT 1920  
Db 1861 CCGATTAAACAGAACCTTGACTTGTATAGCATCTTCTGCTCAGACAGTGTCCGTAT 1920  
Qy 1921 TTGCGAAAAATTCGGGTGATGTCAAAGGATCTTCCAAAGATGATCGGCTGAGCA 1980  
Db 1921 TTGCGAAAAATTCGGGTGATGTCAAAGGATCTTCCAAAGATGATCGGCTGAGCA 1980  
Qy 1981 TGTCTGTGTGTCCTCTATAGATGTGGGATTAAGACCTTCATTTGAGAGAGAGTGGAAAT 2040  
Db 1981 TGTCTGTGTGTCCTCTATAGATGTGGGATTAAGACCTTCATTTGAGAGAGAGTGGAAAT 2040  
Qy 2041 TTCTGCTCATGCTTGCGACCCAGCCACACTGCGGAGCTGGAAGCTGGGAGAGCATC 2100  
Db 2041 TTCTGCTCATGCTTGCGACCCAGCCACACTGCGGAGCTGGAAGCTGGGAGAGCATC 2100  
Qy 2101 CTGACAGAGCGGGCCATGAGAACCTGTGTGCAAGCTGAGGATCCCACTGCAAGATA 2160  
Db 2101 CTGACAGAGCGGGCCATGAGAACCTGTGTGCAAGCTGAGGATCCCACTGCAAGATA 2160  
Qy 2161 CAGACCTGATTTTGAAGATGACAGATTAACCTGCTGTGAGACACTCTGAGAAATC 2220  
Db 2161 CAGACCTGATTTTGAAGATGACAGATTAACCTGCTGTGAGACACTCTGAGAAATC 2220  
Qy 2221 GTCATGGCCAAACCGTAACCTAGATCCCTCAACTTGGAGAGCAACCACTGAAGAAAG 2280  
Db 2221 GTCATGGCCAAACCGTAACCTAGATCCCTCAACTTGGAGAGCAACCACTGAAGAAAG 2280  
Qy 2281 GATGTAAAGATGCGCTGTGAAGCTTAAACACCCAAATGTTTGTGAGTCTTGAAG 2340  
Db 2281 GATGTAAAGATGCGCTGTGAAGCTTAAACACCCAAATGTTTGTGAGTCTTGAAG 2340  
Qy 2341 CTGATTTGCTGTGATGATGACCAATGCTGTTACTGAAAGATCTCCAAATCTTACGAC 2400  
Db 2341 CTGATTTGCTGTGATGATGACCAATGCTGTTACTGAAAGATCTCCAAATCTTACGAC 2400  
Qy 2401 TCCCCAGCTGAATCTCTGAGCCGTGAGAGAAACAAGTGAAGACAGACAGGAGTAATG 2460  
Db 2401 TCCCCAGCTGAATCTCTGAGCCGTGAGAGAAACAAGTGAAGACAGACAGGAGTAATG 2460  
Qy 2461 CCTCTCAGTATGCTTGAAGATCTCCAGTGTGCGCTGCAAGACTGATATCTGAGAGAC 2520  
Db 2461 CCTCTCAGTATGCTTGAAGATCTCCAGTGTGCGCTGCAAGACTGATATCTGAGAGAC 2520  
Qy 2521 TGTGGCATCAAGCAAGCGGTTGCGAGAGTCTGAGCTGAGCCCTGTCAGCAACCGGAGC 2580  
Db 2521 TGTGGCATCAAGCAAGCGGTTGCGAGAGTCTGAGCTGAGCCCTGTCAGCAACCGGAGC 2580  
Qy 2581 TTGAGACACCTGTGCTATCCAAACAAGCTGCGGAAACGAAGGTGAATATCTATGTGT 2640  
Db 2581 TTGAGACACCTGTGCTATCCAAACAAGCTGCGGAAACGAAGGTGAATATCTATGTGT 2640  
Qy 2641 CGATCATGAGGCTTCCCACTGTAGTCTGAGAGAGGCTGATCTGATCAAGTCCACTG 2700  
Db 2641 CGATCATGAGGCTTCCCACTGTAGTCTGAGAGAGGCTGATCTGATCAAGTCCACTG 2700  
Qy 2701 GACAGCGCTGTGTGTTTTCTTGCACCTTGCGCTTAATGAGTAATGCTGACGAC 2760  
Db 2701 GACAGCGCTGTGTGTTTTCTTGCACCTTGCGCTTAATGAGTAATGCTGACGAC 2760  
Qy 2761 CTGACCTTGAAGATGAAACCTGTGGAAGCAATGCGTGAAGCTTGTGCGAGGTCATG 2820  
Db 2761 CTGACCTTGAAGATGAAACCTGTGGAAGCAATGCGTGAAGCTTGTGCGAGGTCATG 2820  
Qy 2821 AGAGAACCATTTGTCATCTCCAGAGACCTGAGTGTGTAAGTGCATCTCACCGCGG 2880  
Db 2821 AGAGAACCATTTGTCATCTCCAGAGACCTGAGTGTGTAAGTGCATCTCACCGCGG 2880

Qy 2881 TGTGTGAGAGTCTGTCTGTGTGATCTGAGAGCAGACACTGGAAGAGCTGTGATCTC 2940  
Db 2881 TGTGTGAGAGTCTGTCTGTGTGATCTGAGAGCAGACACTGGAAGAGCTGTGATCTC 2940  
Qy 2941 ACGAGCAATGCGCTGTGAGAGGCTGTGAGGCTGTGCGCTGTGCGAGAGGACTGAAGCA 3000  
Db 2941 ACGAGCAATGCGCTGTGAGAGGCTGTGAGGCTGTGCGCTGTGCGAGAGGACTGAAGCA 3000  
Qy 3001 AACAGTGTTCGAGAGACTGCGGTTGAAGGATGTGATCTGATCTTGTGATCTGTGAG 3060  
Db 3001 AACAGTGTTCGAGAGACTGCGGTTGAAGGATGTGATCTGATCTTGTGATCTGTGAG 3060  
Qy 3061 GCACTGTGCGGCTTTCGCGCAACCGGATGTGACAGCTTAACCTGCTGAGAAAT 3120  
Db 3061 GCACTGTGCGGCTTTCGCGCAACCGGATGTGACAGCTTAACCTGCTGAGAAAT 3120  
Qy 3121 AACTTCAGTCCCAAGAGATGATGAGCTGTGTTGCGCTTGTGCTGTCCACGCTAAC 3180  
Db 3121 AACTTCAGTCCCAAGAGATGATGAGCTGTGTTGCGCTTGTGCTGTCCACGCTAAC 3180  
Qy 3181 TTACAGATTAATTTGCGCTGTGAAATGCGATGACCTGTGCAATTAAGAGTGTCTGAG 3240  
Db 3181 TTACAGATTAATTTGCGCTGTGAAATGCGATGACCTGTGCAATTAAGAGTGTCTGAG 3240  
Qy 3241 GAAGTGAGCTTACTCAACCCCGAGTCTGTAATGACGCTAATTTGGCATTTCTTTGATAA 3300  
Db 3241 GAAGTGAGCTTACTCAACCCCGAGTCTGTAATGACGCTAATTTGGCATTTCTTTGATAA 3300  
Qy 3301 GATGACCGACCAAAATGAGACTTACTTCCGCGCTCCCTGAAGAGCGGAGCATGCTATG 3360  
Db 3301 GATGACCGACCAAAATGAGACTTACTTCCGCGCTCCCTGAAGAGCGGAGCATGCTATG 3360  
Qy 3361 GCTTGTGCGGAGTGAACCCAGAGCAAGAAAGCTGTGTGCTTCTGCTGTGAGAC 3420  
Db 3361 GCTTGTGCGGAGTGAACCCAGAGCAAGAAAGCTGTGTGCTTCTGCTGTGAGAC 3420  
Qy 3421 TTCAAGAGCATGACCAATTTGCCAAGTCTCTGCTGTGCGCAAGGCAAAATGTCAGTCC 3480  
Db 3421 TTCAAGAGCATGACCAATTTGCCAAGTCTCTGCTGTGCGCAAGGCAAAATGTCAGTCC 3480  
Qy 3481 CAGAGAGTTGAACAAGTGAAGAGCTCCCGCAACCATGAGGAGAGCAAGCAACA 3540  
Db 3481 CAGAGAGTTGAACAAGTGAAGAGCTCCCGCAACCATGAGGAGAGCAAGCAACA 3540  
Qy 3541 CAAGATAAATGTTGAGTGTGATATTCGAGAGCTGTGTGAATCTGTGAGCTGAA 3600  
Db 3541 CAAGATAAATGTTGAGTGTGATATTCGAGAGCTGTGTGAATCTGTGAGCTGAA 3600  
Qy 3601 GGGCTTGAATCAACAGTGTGATCAATGACCAAGAGGTAATGCTGTGCTACATGAGAGA 3660  
Db 3601 GGGCTTGAATCAACAGTGTGATCAATGACCAAGAGGTAATGCTGTGCTACATGAGAGA 3660  
Qy 3661 GAGCTGAGCTGAGAGGCTGTGTGCAACAGTGTGATGACCAAGAGGCTGTGTGCTG 3720  
Db 3661 GAGCTGAGCTGAGAGGCTGTGTGCAACAGTGTGATGACCAAGAGGCTGTGTGCTG 3720  
Qy 3721 CACTGGAGCGGCTGTGAGGCTGTGAGGCTGTGATCAACAGTGTGATGACCAAGCGGT 3780  
Db 3721 CACTGGAGCGGCTGTGAGGCTGTGAGGCTGTGATCAACAGTGTGATGACCAAGCGGT 3780  
Qy 3781 GTGTCTGTGATCACTGGAGCGGCTGTGAGGCTGTGAGGCTGTGATCAACAGTGTGAT 3840  
Db 3781 GTGTCTGTGATCACTGGAGCGGCTGTGAGGCTGTGAGGCTGTGATCAACAGTGTGAT 3840  
Qy 3841 CACAGCGGTGTGAGCTGTGATCACTGGAGCGGCTGTGAGGCTGTGATCAACAGT 3900  
Db 3841 CACAGCGGTGTGAGCTGTGATCACTGGAGCGGCTGTGAGGCTGTGATCAACAGT 3900  
Qy 3901 GCTGATGACCAAGCGGTGTGCTGTGATCACTGGAGCGGCTGTGAGGCTGTGAG 3960  
Db 3901 GCTGATGACCAAGCGGTGTGCTGTGATCACTGGAGCGGCTGTGAGGCTGTGAG 3960



Db 1081 CAAGGGTTCGTCGATCATGAACAACCGTAGCTGCTGACCAAGTGGCCAGGTCGCCGCC 1140  
Qy 1141 GTGGGCTCTCTCATCTGCTGGCCCTGCAAGCTGCAAGAGTGTGGGGGAGAGCTGCGC 1200  
Db 1141 GTGGGCTCTCTCATCTGCTGGCCCTGCAAGCTGCAAGAGTGTGGGGGAGAGCTGCGC 1200  
Qy 1201 CCCTTCAACCAACGCTGCAAGGCTGCAAGGCTGCTTTTGTGTTTATCAGCTCAACCCCT 1260  
Db 1201 CCCTTCAACCAACGCTGCAAGGCTGCAAGGCTGCTTTTGTGTTTATCAGCTCAACCCCT 1260  
Qy 1261 CGAGGCGTGGTCCGGGCTGCTCTCATCTGAGAGAAAGATTGCTCTGAAGGCTTCTG 1320  
Db 1261 CGAGGCGTGGTCCGGGCTGCTCTCATCTGAGAGAAAGATTGCTCTGAAGGCTTCTG 1320  
Qy 1321 CGTATGGCTGTGAGAGGAGTGTGAATAGAAAGTCACTGTTGATGTGAACAAGCTCATG 1380  
Db 1321 CGTATGGCTGTGAGAGGAGTGTGAATAGAAAGTCACTGTTGATGTGAACAAGCTCATG 1380  
Qy 1381 GTTCAAGGACTGGGGGAGTGTGAGCTCGGTGCTGTTCATCAATGAACATCCCTTCCCA 1440  
Db 1381 GTTCAAGGACTGGGGGAGTGTGAGCTCGGTGCTGTTCATCAATGAACATCCCTTCCCA 1440  
Qy 1441 GACAGCCTGTGAGAGAGTACTACACTTCTTCACTGAGTCTCCAGGACTTCTGTGCG 1500  
Db 1441 GACAGCCTGTGAGAGAGTACTACACTTCTTCACTGAGTCTCCAGGACTTCTGTGCG 1500  
Qy 1501 GCGTTGTACTAGTGTGAAGAGGCTGTGAATGAGCCAGCTCTGCGCTCTGTACGTT 1560  
Db 1501 GCGTTGTACTAGTGTGAAGAGGCTGTGAATGAGCCAGCTCTGCGCTCTGTACGTT 1560  
Qy 1561 GAGAGAGCAAGAGGCTGCATGAGGCTTAAACAGGCGGCTTCATATCCACTGCGTTTG 1620  
Db 1561 GAGAGAGCAAGAGGCTGCATGAGGCTTAAACAGGCGGCTTCATATCCACTGCGTTTG 1620  
Qy 1621 ATGAAGCGTTTCTGTGTTGGGCTCTGAGCGAAGAGTAAAGAGGCACTGAGAGTCTG 1680  
Db 1621 ATGAAGCGTTTCTGTGTTGGGCTCTGAGCGAAGAGTAAAGAGGCACTGAGAGTCTG 1680  
Qy 1681 CTGGGCTGTCCGTTCCCTGGGGGTGAAGCAGAGCTTCTGCACTGGTCTCTCTGTG 1740  
Db 1681 CTGGGCTGTCCGTTCCCTGGGGGTGAAGCAGAGCTTCTGCACTGGTCTCTCTGTG 1740  
Qy 1741 GGTGAGGAGCTTAAAGGCAACCCGAGAGAGCACTGAGAGGCTTCCACTGCTTTTC 1800  
Db 1741 GGTGAGGAGCTTAAAGGCAACCCGAGAGAGCACTGAGAGGCTTCCACTGCTTTTC 1800  
Qy 1801 GAGACTCAAGACAAAGAGTTTGTGCTTGGCATTTAAACAGCTTCCAAAGAGTGTGCT 1860  
Db 1801 GAGACTCAAGACAAAGAGTTTGTGCTTGGCATTTAAACAGCTTCCAAAGAGTGTGCT 1860  
Qy 1861 CGGATTTAACAGAACCTGAGCTTGAATGATCTTCTGCTGCTCCAGAGCTGTCCGAT 1920  
Db 1861 CGGATTTAACAGAACCTGAGCTTGAATGATCTTCTGCTGCTCCAGAGCTGTCCGAT 1920  
Qy 1921 TTGGCGGAAAATTCCGGGTGATGTCAAAAGGAGTCTTCCCAAGAGTAAAGTCCCTGAGGA 1980  
Db 1921 TTGGCGGAAAATTCCGGGTGATGTCAAAAGGAGTCTTCCCAAGAGTAAAGTCCCTGAGGA 1980  
Qy 1981 TGTCTGTGGTCCCTCTATGATGTGGAGTAAAGACCTCATTTAGAGAGAGTGTGAGAAAT 2040  
Db 1981 TGTCTGTGGTCCCTCTATGATGTGGAGTAAAGACCTCATTTAGAGAGAGTGTGAGAAAT 2040  
Qy 2041 TTCTGCTCATGCTTGGCAACCAACCACTGCGGAGCTGAGACTTGGGCAAGCAGCATC 2100  
Db 2041 TTCTGCTCATGCTTGGCAACCAACCACTGCGGAGCTGAGACTTGGGCAAGCAGCATC 2100  
Qy 2101 CTGAGCAGAGCGGCGCATGAAGACCTGTGTGCCAAGCTGAGGATCCCACTGCAAGATA 2160  
Db 2101 CTGAGCAGAGCGGCGCATGAAGACCTGTGTGCCAAGCTGAGGATCCCACTGCAAGATA 2160  
Qy 2161 CAGACCTGATGTTTGAATGACAGATTAACCTGCTGTGTGAGAGCACTCTGAGAAATC 2220  
Db 2161 CAGACCTGATGTTTGAATGACAGATTAACCTGCTGTGTGAGAGCACTCTGAGAAATC 2220

Qy 2221 GTCATGGCAACCGTAACCTTAAGATCCCTCAACTTGGAGGAGCACCTGAAGAGAG 2280  
Db 2221 GTCATGGCAACCGTAACCTTAAGATCCCTCAACTTGGAGGAGCACCTGAAGAGAG 2280  
Qy 2281 GATGTAAAGATGGCTGTGAAGCTTAAACACCCAAATGTTTGTGAGTCTTTGAGG 2340  
Db 2281 GATGTAAAGATGGCTGTGAAGCTTAAACACCCAAATGTTTGTGAGTCTTTGAGG 2340  
Qy 2341 CTGATGTGCTGTGATGAACCATGCTGTAACTGAAGATCTCCAAATCTTAAGAC 2400  
Db 2341 CTGATGTGCTGTGATGAACCATGCTGTAACTGAAGATCTCCAAATCTTAAGAC 2400  
Qy 2401 TCCCGGAGCTGAAATCTCTGAGCTGTGAGAGAAACAGGTGACAGACCAAGGAAATATG 2460  
Db 2401 TCCCGGAGCTGAAATCTCTGAGCTGTGAGAGAAACAGGTGACAGACCAAGGAAATATG 2460  
Qy 2461 CCTCTCAGTATGCTTGAAGAGTCTCCAGTACGCGCTGAGAGCTGATCTGAGAGAC 2520  
Db 2461 CCTCTCAGTATGCTTGAAGAGTCTCCAGTACGCGCTGAGAGCTGATCTGAGAGAC 2520  
Qy 2521 TGTGGCATCAAGCCAGGTTTCCAGAGTCTGAGCTCAAGCTTGTGAGCAACCGGAGC 2580  
Db 2521 TGTGGCATCAAGCCAGGTTTCCAGAGTCTGAGCTCAAGCTTGTGAGCAACCGGAGC 2580  
Qy 2581 TTGACACACTGTGCTATCCCAACACAGCTGTGGGAAACGAAGTGTAAATCTACTGT 2640  
Db 2581 TTGACACACTGTGCTATCCCAACACAGCTGTGGGAAACGAAGTGTAAATCTACTGT 2640  
Qy 2641 CGATCCATGAGGCTTCCCACTGTAGTGTGAGAGGCTGATGCTGAATCAGTGCCACTG 2700  
Db 2641 CGATCCATGAGGCTTCCCACTGTAGTGTGAGAGGCTGATGCTGAATCAGTGCCACTG 2700  
Qy 2701 GACAGGCTGTGCTGTTTTCTTTCACCTTGCTTAATGAGTAACTCATGCTGACGAC 2760  
Db 2701 GACAGGCTGTGCTGTTTTCTTTCACCTTGCTTAATGAGTAACTCATGCTGACGAC 2760  
Qy 2761 CTGAGCTTAAAGTAAACCTGTGTGAAGACATGAGCTGTGAGAGTCTGTGAGGTCATG 2820  
Db 2761 CTGAGCTTAAAGTAAACCTGTGTGAAGACATGAGCTGTGAGAGTCTGTGAGGTCATG 2820  
Qy 2821 AGAGAACCATCTTGCATCTCCAGAGCTGAGAGTGTAAAGTGCATCTCACGCGCG 2880  
Db 2821 AGAGAACCATCTTGCATCTCCAGAGCTGAGAGTGTAAAGTGCATCTCACGCGCG 2880  
Qy 2881 TGTGTGAGAGTCTGTCTGTGTGATCTGAGAGACAGACACTGAAAGCTGTGATCTC 2940  
Db 2881 TGTGTGAGAGTCTGTCTGTGTGATCTGAGAGACAGACACTGAAAGCTGTGATCTC 2940  
Qy 2941 ACGACAAATGCCCTGGGTGAAGGTGGGTGTGCTGCTGTGAGAGGACTGAAGCAAAAG 3000  
Db 2941 ACGACAAATGCCCTGGGTGAAGGTGGGTGTGCTGCTGTGAGAGGACTGAAGCAAAAG 3000  
Qy 3001 AACAGTGTGAGAGAGCTGGGGTTGAAGGAGTGAAGTGAAGTCTGATGCTGTGAG 3060  
Db 3001 AACAGTGTGAGAGAGCTGGGGTTGAAGGAGTGAAGTGAAGTCTGATGCTGTGAG 3060  
Qy 3061 GCACTCTCTTGGCCCTTCTGCAACCGGATCTGACAGTCTAAACCTGTGTGAGAT 3120  
Db 3061 GCACTCTCTTGGCCCTTCTGCAACCGGATCTGACAGTCTAAACCTGTGTGAGAT 3120  
Qy 3121 AACTTCAGTCCCAAGAGATGATGAAGCTGTGTGCTTGGCTTGTCCCAAGCTTAAC 3180  
Db 3121 AACTTCAGTCCCAAGAGATGATGAAGCTGTGTGCTTGGCTTGTCCCAAGCTTAAC 3180  
Qy 3181 TTACAGATTAATTTGGCTGTGAAATGCAAGTAACTGTGCAAAATPAAGAGTGTGAG 3240  
Db 3181 TTACAGATTAATTTGGCTGTGAAATGCAAGTAACTGTGCAAAATPAAGAGTGTGAG 3240  
Qy 3241 GAAAGTGAAGTAACTCAAGCCCGAGTCTGAATTTGACGATGTGAGCTTTTGTGATGA 3300  
Db 3241 GAAAGTGAAGTAACTCAAGCCCGAGTCTGAATTTGACGATGTGAGCTTTTGTGATGA 3300

QY	3301	TATGACCGACACAAATATGACTTACTTTCCGGCTCCCTGAAAGCCGGGCAATGCAATG	3366
Db	3301	GATGACCGACACAAATATGACTTACTTTCCGGCTCCCTGAAAGCCGGGCAATGCAATG	3366
QY	3361	GCTTGCTGTGGGGGATGAACCCAGAGCAGAAAGACTGTGTGCTTCTTGAGTGGAGAC	3420
Db	3361	GCTTGCTGTGGGGGATGAACCCAGAGCAGAAAGACTGTGTGCTTCTTGAGTGGAGAC	3420
QY	3421	TTCAAGACGAGTACACGATTTTGCCAAAGTCTCTGTGCTGGCCACCGGCAATGTGTGTC	3480
Db	3421	TTCAAGACGAGTACACGATTTTGCCAAAGTCTCTGTGCTGGCCACCGGCAATGTGTGTC	3480
QY	3481	CAGAGAGTTGACAACTGTGAGAGAGGCTCCCGCAACCCATGTGAGGAGACGGAAACAA	3544
Db	3481	CAGAGAGTTGACAACTGTGAGAGAGGCTCCCGCAACCCATGTGAGGAGAGGAAACAA	3544
QY	3541	CAAGATATAATGTGAGTGTGGAATATTCGGAGGCTGCTGTCTGAAACTGTGAGCTCGAA	3600
Db	3541	CAAGATATAATGTGAGTGTGGAATATTCGGAGGCTGCTGTCTGAAACTGTGAGCTCGAA	3600
QY	3601	GGGCTTGATCCAAACAGTGTGATCATGACCAACGAGAGTATGCGCTGTGTCACTAGGAGA	3666
Db	3601	GGGCTTGATCCAAACAGTGTGATCATGACCAACGAGAGTATGCGCTGTGTCACTAGGAGA	3666
QY	3661	GAGCTGAGCTCGAGGGGCTTGTGTCTCAACAGTGTGATGACCAACGCGGTGTCTTGT	3722
Db	3661	GAGCTGAGCTCGAGGGGCTTGTGTCTCAACAGTGTGATGACCAACGCGGTGTCTTGT	3722
QY	3721	CACGTGGAGCGGCTGGGCTCTAGGGGGCGGTGTCTTAACAGTGTGATGACCAACAGCGGT	3780
Db	3721	CACGTGGAGCGGCTGGGCTCTAGGGGGCGGTGTCTTAACAGTGTGATGACCAACAGCGGT	3780
QY	3781	GTGTCTGTGTCACTGTGGAGCGGCTGTGGCTCGAGGGGCTTGTGTCCAAACAGTGTGATGAC	3840
Db	3781	GTGTCTGTGTCACTGTGGAGCGGCTGTGGCTCGAGGGGCTTGTGTCCAAACAGTGTGATGAC	3840
QY	3841	CACAGCGGTGTGGCTTGTGTCACTGTGGAGCGGCTGTGGCTCGAGGGGCTTGTGTCCAAACAGT	3900
Db	3841	CACAGCGGTGTGGCTTGTGTCACTGTGGAGCGGCTGTGGCTCGAGGGGCTTGTGTCCAAACAGT	3900
QY	3901	GCTGATGACCAAGCGGTGTGCTGTGTCACTGTGGAGGGGCTGTGGGCTGTGAGAGGGGCTTGTG	3966
Db	3901	GCTGATGACCAAGCGGTGTGCTGTGTCACTGTGGAGGGGCTGTGGGCTGTGAGAGGGGCTTGTG	3966
QY	3961	TCCAACAGTGTGTGATGACCAACGCGGTGTGTCTTGTCACTGTGGAGCGGCTGTGGGCTCGAG	4020
Db	3961	TCCAACAGTGTGTGATGACCAACGCGGTGTGTCTTGTCACTGTGGAGCGGCTGTGGGCTCGAG	4020
QY	4021	GGGCTGTGTGTCTTAA 4035	
Db	4021	GGGCTGTGTGTCTTAA 4035	

	RESULT 3
ABK48628	
ID	ABK48628 standard; cDNA; 3900 BP.
XX	
AC	ABK48628;
XX	
DT	13-AUG-2002 (first entry)
XX	
DE	Human MATER cDNA.
XX	
KW	Human; gene; ss; contraceptive; antiinfertility; MATER;
KW	maternal antigen that embryos require; MATER null phenotype; oocyte
KW	early embryonic survival; premature ovarian failure; POB;
KW	autoimmune infertility; chromosome 19; gene therapy; fertility.
XX	
OS	Homo sapiens.
XX	
XX	
Key	Location/Qualifiers
FT	1..3603
FT	/*tag= a
FT	

FT	/product= "Human MATER"
XX	
FN	WO200232955-A1.
PD	25-APR-2002.
XX	
PF	04-APR-2001; 2001MO-US010981.
XX	
PR	18-OCT-2000; 2000US-0241510P.
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI	Nelson LM, Tong Z;
XX	
DR	MP1: 2002-454595/48.
XX	P-PsDB; AAU79526.
PT	New isolated human Maternal Antigen That Embryos Require protein and polynucleotide, useful in diagnosing or treating fertility or reduced fertility, or as a contraceptive.
PS	Claim 11; Page 80-85; 93pp; English.
XX	
CC	The invention discloses an isolated human MATER (Maternal Antigen That Embryos Require) protein, which can complement a MATER null phenotype in which zygotes arising from the oocyte do not progress beyond the two-cell stage. MATER is required for early embryonic survival and abnormal levels of the protein can lead to premature ovarian failure (POF) and can be caused by under or over expression of MATER or an autoimmune response to MATER. MATER is a single-copy maternal effect gene found on chromosome 19. The MATER protein and polynucleotide, by gene therapy, are useful in diagnosing or treating fertility and reduced fertility. In particular, the MATER protein is useful as a contraceptive agent, or for influencing (either inhibiting or enhancing) fertility and can be used to detect a predisposition to infertility or reduced fertility, or for presymptomatic screening of an individual for infertility/reduced fertility. The protein and polynucleotide are also useful for detecting an excess or deficiency, or genetic mutation, of the MATER protein in a mammalian subject (e.g. a human or a mouse) or for screening for a compound useful in influencing MATER-mediated fertility. The sequence presented is the human MATER cDNA. Note: Nucleotides 1-75 were deduced, by the inventors, by comparison to published human genomic DNA sequences, while the remainder of the sequence was determined by direct cloning of human ovarian CDNA's
SQ	Sequence 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;
Query Match	78.5%; Score 3166.2; DB 6; Length 3900;
Best Local Similarity	96.2%; Pred. No. 0;
Matches 3304; Conservative	0; Mismatches 8; Indels 123; Gaps 11
OY	1 ATGGAAGAGACAATCGCTCAACCTTTTCAGCTACGGGCTGCATGTGTCCTTAGAG 60     DB 154 ATGGAAGAGACAATCGCTCAACCTTTTCAGCTACGGGCTGCATGTGTCCTTAGAG 213 
OY	61 CTAGCAAGAAGAAATTTTCAAGCATTCGAAGAAATTACTAAGAAAGAAATCTTGCAATCG 120     DB 214 CTAGCAAGAAGAAATTTTCAAGCATTCGAAGAAATTACTAAGAAAGAAATCTTGCAATCG 273 
OY	121 ACCACATCTCTAATTCACAGTTTGAATCGAATGCAACGTGGAATGTCGTGGCATTTC 180     DB 274 ACCACATCTCTAATTCACAGTTTGAATGCAATGCAACGTGGAATGTCGTGGCATTTC 333 
OY	181 CTCCTGATGAGTATTAATGAGCATCGCTGGCTCTGGCTACGTCATTAGCATCTTTGAA 240     DB 334 CTCCTGATGAGTATTAATGAGCATCGCTGGCTCTGGCTACGTCATTAGCATCTTTGAA 393 
OY	241 AACATGAACCTGGCAACCTCTCGGAGAAAGGCAAGGATGACATGAAAA----- 289     DB 394 AACATGAACCTGGCAACCTCTCGGAGAAAGGCAAGGATGACATGAAAAACATTACCA 453 
OY	290 ----- 289
DB	454 GAAGTCTCTGAGCAAGATGACTGACCAGAACCAAGCAAGAAAAGTGCAGGAATT 513

290 -----AAATTGA 297  
Db 514 TCACAACTGTGCAACAATAGTGTGCAACGCTGCAGAGACAAAGAACAGAAATTTCACAA 573  
Qy 298 CAAGCTATGGAACAAGAGGTGCAACAGACAGAGACAGAAACAAAGAAATTTTCACAA 357  
Db 574 CAAGCTATGGAACAAGAGGTGCAACAGACAGAGACAGAAACAAAGAAATTTTCACAA 633  
Qy 358 GCTATGGAACAAGAGGTGCAACAGACAGAGACAGAAACAAAGAAATTTTCACAA 417  
Db 634 GCTATGGAACAAGAGGTGCAACAGACAGAGACAGAAACAAAGAAATTTTCACAA 693  
Qy 418 ACATGGGATCAACAGAGTCAACGATGATGACCAATTGCTGAGAGAGAGATGATGATGAT 477  
Db 694 ACATGGGATCAACAGAGTCAACGATGATGACCAATTGCTGAGAGAGAGATGATGATGAT 753  
Qy 478 AGTTTGAACACATGCTGCTGACCTGACCGAAATGCAACGTTGCTGCTGCTTTGAT 537  
Db 754 AGTTTGAACACATGCTGCTGACCTGACCGAAATGCAACGTTGCTGCTGCTTTGAT 813  
Qy 538 TCAGACCGGTGGGCTTCCGCTGCAACGCTGCTGCAACGAAAGTCAAGAAATTGGG 597  
Db 814 TCAGACCGGTGGGCTTCCGCTGCAACGCTGCTGCAACGAAAGTCAAGAAATTGGG 873  
Qy 598 AATGCGCTCTAGCCAGAGATCGTGTGCTGGGCGCAAGTGGATCTTACCAAGGA 657  
Db 874 AATGCGCTCTAGCCAGAGATCGTGTGCTGGGCGCAAGTGGATCTTACCAAGGA 933  
Qy 658 ATGTTCTCTTACGCTCTTCTCTCCCGTTAGAGAGATGACAGCGGAAGAGAGAGACAT 717  
Db 934 ATGTTCTCTTACGCTCTTCTCTCCCGTTAGAGAGATGACAGCGGAAGAGAGAGACAT 993  
Qy 718 GTCAAGAGTTCATCTTCAAGAGATGAGGCAAGCTCCAGGCTCCGCTGACAGAGATCAT 777  
Db 994 GTCAAGAGTTCATCTTCAAGAGATGAGGCAAGCTCCAGGCTCCGCTGACAGAGATCAT 1053  
Qy 778 TCCCGACCAAGAGAGTGTGTTTTCATCATTAAGGTTTGAATGACCTGGGCTCTGCTTC 837  
Db 1054 TCCCGACCAAGAGAGTGTGTTTTCATCATTAAGGTTTGAATGACCTGGGCTCTGCTTC 1113  
Qy 838 AACATGACACAAAGCTCTGCAAGAGCTGGGCTGAGAACAGCTCCGTTACCCCTCAT 897  
Db 1114 AACATGACACAAAGCTCTGCAAGAGCTGGGCTGAGAACAGCTCCGTTACCCCTCAT 1173  
Qy 898 GCGAGTCTGTGAGAGAGGTCCTGCTCCCTGAGTCTCTTCTGATGTCACCGTCAAGAG 957  
Db 1174 GCGAGTCTGTGAGAGAGGTCCTGCTCCCTGAGTCTCTTCTGATGTCACCGTCAAGAG 1233  
Qy 958 GTGGGCAACAGAGAGTCAAGTCAAGAGTCTGCTGCTCCCGTTACCTGTTAGTGAAGA 1017  
Db 1234 GTGGGCAACAGAGAGTCAAGTCAAGAGTCTGCTGCTCCCGTTACCTGTTAGTGAAGA 1293  
Qy 1018 ATCTCGGGGGAACAAGAAATCCACTTGTCTTGAAGCGGGATTGGTGAATCAGAAAG 1077  
Db 1294 ATCTCGGGGGAACAAGAAATCCACTTGTCTTGAAGCGGGATTGGTGAATCAGAAAG 1353  
Qy 1078 ACACAAGAGGTTGCTGCTGATGATGACAAACGCTGAGCTGCTGACCAAGTCCAGGTGCC 1137  
Db 1354 ACACAAGAGGTTGCTGCTGATGATGACAAACGCTGAGCTGCTGACCAAGTCCAGGTGCC 1413  
Qy 1138 GCGGTGGGCTCTTCTCATCTGCTGCTGCTGCAAGCTGCAAGAGCTGGTGGGAGAGAGCT 1197  
Db 1414 GCGGTGGGCTCTTCTCATCTGCTGCTGCTGCAAGCTGCAAGAGCTGGTGGGAGAGAGCT 1473  
Qy 1198 GCGGTGGGCTCTTCTCATCTGCTGCTGCTGCAAGCTGCAAGAGCTGGTGGGAGAGAGCT 1257  
Db 1474 GCGGTGGGCTCTTCTCATCTGCTGCTGCTGCAAGCTGCAAGAGCTGGTGGGAGAGAGCT 1533  
Qy 1258 CCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1317  
Db 1534 CCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1593

Qy 1318 TGCCGTATGCTGTGAGAGAGTGTGAATAGAGATCAGTGTGTTAGTGTGACAGACTTC 1377  
Db 1594 TGCCGTATGCTGTGAGAGAGTGTGAATAGAGATCAGTGTGTTAGTGTGACAGACTTC 1653  
Qy 1378 ATGTTCAAGGACTCGGGGAGTCTGAGCTCCGCTGCTGTTTCAATGAACATCTTCTC 1437  
Db 1654 ATGTTCAAGGACTCGGGGAGTCTGAGCTCCGCTGCTGTTTCAATGAACATCTTCTC 1713  
Qy 1438 CCAGACCCACTGTGAGAGAGTCAACACTTCTTCCACCTCAAGTCTCAGAGACTTCTGT 1497  
Db 1714 CCAGACCCACTGTGAGAGAGTCAACACTTCTTCCACCTCAAGTCTCAGAGACTTCTGT 1773  
Qy 1498 GCGGCTGTACTAGTGTGAGAGGCTGGAATTCAGACAGCTCTGCTGCTGCTGCTGCT 1557  
Db 1774 GCGGCTGTACTAGTGTGAGAGGCTGGAATTCAGACAGCTCTGCTGCTGCTGCTGCTG 1833  
Qy 1558 GTTGAAGAGACAAAGAGTCCATGAGAGCTTAAACAGGAGCTTCCATATCCACTCGCT 1617  
Db 1834 GTTGAAGAGACAAAGAGTCCATGAGAGCTTAAACAGGAGCTTCCATATCCACTCGCT 1893  
Qy 1618 TGATGAAGCGTTTCTTGTGCTGCTGCTGAGCCGAAGCTTAAGAGAGCTCAAGAGTCT 1677  
Db 1894 TGATGAAGCGTTTCTTGTGCTGCTGCTGAGCCGAAGCTTAAGAGAGCTCAAGAGTCT 1953  
Qy 1678 CTGCTGGGCTGCTCCGCTCCCGTGGGGTGAAGAGAGCTTCTGCTGCTGCTGCTGCT 1737  
Db 1954 CTGCTGGGCTGCTCCGCTCCCGTGGGGTGAAGAGAGCTTCTGCTGCTGCTGCTGCTG 2013  
Qy 1738 TTGGGTGAGCAGCTTAAATGCAACCAACCCAGAGAGACCTTGAAGCTTCACTGCTCT 1797  
Db 2014 TTGGGTGAGCAGCTTAAATGCAACCAACCCAGAGAGACCTTGAAGCTTCACTGCTCT 2073  
Qy 1798 TTGAGACTCAAGACAAAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1857  
Db 2074 TTGAGACTCAAGACAAAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2133  
Qy 1858 CTTCGATTAACCAAGACCTGAGCTGATGATCTTCTGCTGCTGCTGCTGCTGCTGCTG 1917  
Db 2134 CTTCGATTAACCAAGACCTGAGCTGATGATCTTCTGCTGCTGCTGCTGCTGCTGCTG 2193  
Qy 1918 TATTTGCGGAAATTCGGGTGATGTCAAAAGGATCTTCCCAAGAGATGATGCTGCTG 1977  
Db 2194 TATTTGCGGAAATTCGGGTGATGTCAAAAGGATCTTCCCAAGAGATGATGCTGCTG 2253  
Qy 1978 GCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2037  
Db 2254 GCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2313  
Qy 2038 GATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2097  
Db 2314 GATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2373  
Qy 2098 ATCTGACAGAGCGGGCATGAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2157  
Db 2374 ATCTGACAGAGCGGGCATGAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2433  
Qy 2158 ATACAGACCTGATGTTTGAAGATGACAGATTAACCTGCTGCTGCTGCTGCTGCTG 2217  
Db 2434 ATACAGACCTGATGTTTGAAGATGACAGATTAACCTGCTGCTGCTGCTGCTGCTG 2493  
Qy 2218 ATGCTATGAGGCAACCTGATTAAGATTCCTCAACTTGGAGGCAACCACTGAAGAA 2277  
Db 2494 ATGCTATGAGGCAACCTGATTAAGATTCCTCAACTTGGAGGCAACCACTGAAGAA 2553  
Qy 2278 GAGGATGAAGAGTGGCTGTAAGCTTAAACACCAAAATGTTTGTGAGAGCTTGG 2337  
Db 2554 GAGGATGAAGAGTGGCTGTAAGCTTAAACACCAAAATGTTTGTGAGAGCTTGG 2613  
Qy 2338 AGGCTGATGCTGCTGATGATGACCAATGCTGTTAATCTGGAAGATCTCCCAATCTTAC 2397  
Db 2614 AGGCTGATGCTGCTGATGATGACCAATGCTGTTAATCTGGAAGATCTCCCAATCTTAC 2673  
Qy 2398 AACTCCCCAGCTGGAATCTCTGAGCTGGCAGAGAAACAAGGTGACAGACCAAGAGATA 2457

```
Db 2674 ACCTCCCCAGCCTGAATCTCTGAGCCTGGCAGGAAACAAGTGAACAGACACGAGGAGTA 2733
QY 2458 ATGCTCTCAGTGAAGCCTTGAAGAGTCTCCAGTGGCGCCCTGCANAACCTGATCTGAG 2517
Db 2734 ACGCTCTCAGTGAAGCCTTGAAGAGTCTCCAGTGGCGCCCTGCANAACCTGATCTGAG 2793
QY 2518 GACTGTGCAATCAAGCAGCGGTTGCCAGAGTCTGGCCTCAGCCCTGTCAGCAACCGG 2577
Db 2794 GACTGTGCAATCAAGCAGCGGTTGCCAGAGTCTGGCCTCAGCCCTGTCAGCAACCGG 2853
QY 2578 AGCTTGAACAACCTGTGCTTATCCAAACAAGCCTGGGGAACGAAGTGTAAATCTACTG 2637
Db 2854 AGCTTGAACAACCTGTGCTTATCCAAACAAGCCTGGGGAACGAAGTGTAAATCTACTG 2913
QY 2638 TGTGATCAATGAAGCCTCCCACTGTAGTCTGGCAGAGGCTGATCTGAATCAGTCCAC 2697
Db 2914 TGTGATCAATGAAGCCTCCCACTGTAGTCTGGCAGAGGCTGATCTGAATCAGTCCAC 2973
QY 2698 CTGACACGCGCTGCTGTGTTTCTTTCACCTTGGCTTAATGGTAACTCATAGGCTGACG 2757
Db 2974 CTGACACGCGCTGCTGTGTTTCTTTCACCTTGGCTTAATGGTAACTCATAGGCTGACG 3033
QY 2758 CACTGAGCCTTGAATGAACCTCTGTGAAGAACAATGCGGTGAACCTTCTGTGCAAGTTC 2817
Db 3034 CACTGAGCCTTGAATGAACCTCTGTGAAGAACAATGCGGTGAACCTTCTGTGCAAGTTC 3093
QY 2818 ATGAGAAGAACCATCTGTGATCTCCAGACCTGGAGTTGTAAATGTCAATCCAGCC 2877
Db 3094 ATGAGAAGAACCATCTGTGATCTCCAGACCTGGAGTTGTAAATGTCAATCCAGCC 3153
QY 2878 GCGTCTGTGAAGTCTGTCTGTGTGATCTTCGAGAGACAGACACTGAAGAGCTGTGAT 2937
Db 3154 GCGTCTGTGAAGTCTGTCTGTGTGATCTTCGAGAGACAGACACTGAAGAGCTGTGAT 3213
QY 2938 CTCACGGAACAATGCTTGGGTGAACGTTGGGTTGCTGCGCTGTGCGAGGACTGAAGCAA 2997
Db 3214 CTCACGGAACAATGCTTGGGTGAACGTTGGGTTGCTGCGCTGTGCGAGGACTGAAGCAA 3273
QY 2998 AAGAACAGTGTCTGAACGAGACTCGGGTTGAAGGAGATGGAATGATCTTGAATGCTGT 3057
Db 3274 AAGAACAGTGTCTGAACGAGACTCGGGTTGAAGGAGATGGAATGATCTTGAATGCTGT 3333
QY 3058 GAGGACTCTCTTGGCCCTTCTGCAACCGGACTGTGAACGATCTAACTGGTGGAG 3117
Db 3334 GAGGACTCTCTTGGCCCTTCTGCAACCGGACTGTGAACGATCTAACTGGTGGAG 3393
QY 3118 AATTAATCTCACTCCCAAGAAATGATGAAGCTGTGCGCTTGTGCTGTCCAGCTCT 3177
Db 3394 AATTAATCTCACTCCCAAGAAATGATGAAGCTGTGCGCTTGTGCTGTCCAGCTCT 3453
QY 3178 AACTTAAGATTAATTTGGGCTGTGGAATGGCAGTACCTGTGCAATAATGAAGAGTGTCTG 3237
Db 3454 AACTTAAGATTAATTTGGGCTGTGGAATGGCAGTACCTGTGCAATAATGAAGAGTGTCTG 3513
QY 3238 GAGGAAGTGAAGTCTCAAGCCCGAGTGTGAATTTGAAGTGTGAGTGTGCAATCTTTGAT 3297
Db 3514 GAGGAAGTGAAGTCTCAAGCCCGAGTGTGAATTTGAAGTGTGAGTGTGCAATCTTTGAT 3573
QY 3298 GAAGATGAACGACAC 3312
Db 3574 GAAGATGAACGAGTAC 3588
```

RESULT 4  
AAD49018  
ID AAD49018 standard; cDNA; 3900 BP.

AC AAD49018;  
XX  
DT 07-MAR-2003 (first entry)  
XX  
DE Human MATER cDNA.

```
XX Human; MATER protein; infertility; fertility; contraceptive agent;  
KW gene therapy; gene; ss.  
XX  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 1..3603  
FT /*tag= a  
FT /product= "Human MATER protein"  
XX  
EN WO200281492-A1.  
PD 17-OCT-2002.  
XX  
PE 29-MAR-2002; 2002WO-US009776.  
XX  
PR 04-APR-2001; 2001WO-US010981.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Nelson LM, Tong Z;  
XX  
DR WPI; 2003-058494/05.  
XX P-PSDB; AAE31749.  
XX  
PT New isolated variant MATER proteins and nucleic acids, useful for  
PT diagnosing, prognosing and treating infertility and reduced fertility,  
PT and as contraceptive agents.  
XX  
PS Example 2; Page 93-98; 110pp; English.  
XX  
CC The present invention relates to novel MATER proteins and polynucleotides  
CC encoding such proteins. The MATER proteins are essential to fertility.  
CC Sequences of the invention are useful for diagnosing, prognosing and  
CC treating infertility, reduced fertility and as contraceptive agents. They  
CC are also useful in gene therapy. The method is useful for detecting a  
CC predisposition to or pre-symptomatic screening of an individual for  
CC infertility or reduced fertility. The present sequence is human MATER  
CC cDNA  
XX  
SQ Sequence 3900 BP; 981 A; 987 C; 1044 G; 888 T; 0 U; 0 Other;  
XX  
Query Match 78.5%; Score 3166.2; DB 8; Length 3900;  
Best Local Similarity 96.2%; Pred. No. 0;  
Matches 3304; Conservative 0; Mismatches 8; Indels 123; Gaps 1;  
QY 1 ATGAGAAGAGCAAAATGCTCACTTTTCCAGCTACGCGGCTGCAATGCTGTCTTATGAG 60  
Db 154 ATGAGAAGAGCAAAATGCTCACTTTTCCAGCTACGCGGCTGCAATGCTGTCTTATGAG 213  
QY 61 CTAAACAAGAAAGAAATTTGAGACATTCAGAGAAATTAATTAAGAAAGAAATCTTCAGAAATCG 120  
Db 214 CTAAACAAGAAAGAAATTTGAGACATTCAGAGAAATTAATTAAGAAAGAAATCTTCAGAAATCG 273  
QY 121 ACCACATGCTTATTTCCAGATTTGAATCGAAGATCCAACTGGAATGTCTGGCACTC 180  
Db 274 ACCACATGCTTATTTCCAGATTTGAATCGAAGATCCAACTGGAATGTCTGGCACTC 333  
QY 181 CTCTTGATGAGATTTATGAGATCGCTGGCTGGCTGACATTCATTAAGATCTTTGAA 240  
Db 334 CTCTTGATGAGATTTATTTATGAGATCGCTGGCTGGCTGACATTCATTAAGATCTTTGAA 393  
QY 241 AACATGAACCTGGAGACCTCTCGGAAGAGCAAGGATGATCATGAATAA----- 289  
Db 394 AACATGAACCTGGAGACCTCTCGGAAGAGCAAGGATGATCATGAATAAAGCATTCACCA 453  
QY 290 ----- 289  
Db 454 GAAGATCTGAAGCAAGATGACTGACCAAGAGCAAGCAAGAAAGTCCAGAAAT 513  
QY 290 -----AAATTTCA 297
```



Db 514 TCACAGCTGTGCACACAGATAGTCCACAGCTGCAGAGACAAAGAAACAGAAATTTCAC 573  
Qy 298 CAAGCTATGGAACAAAGAGGTGCACAGCAGACAGAGAAAGAAATTTTCACAA 357  
Db 574 CAAGCTATGGAACAAAGAGGTGCACAGCAGACAGAGAAAGAAATTTTCACAA 633  
Qy 358 GCTATGGAACAAAGAGGTGCACAGCAGACAGAGAAAGAAATTTTCACAA 417  
Db 634 GCTATGGAACAAAGAGGTGCACAGCAGACAGAGAAAGAAATTTTCACAA 693  
Qy 418 ACATGGAGCTACAGAGTCAAGTATGACCAATTGCTGAGAGAGATGTACGTCT 477  
Db 694 ACATGGAGCTACAGAGTCAAGTATGACCAATTGCTGAGAGAGATGTACGTCT 753  
Qy 478 AATTTTGAACACATGCTGCTGATGCTGCGGAAATTCACAAAGTGTGCTGCTTTGAT 537  
Db 754 AATTTTGAACACATGCTGCTGATGCTGCGGAAATTCACAAAGTGTGCTGCTTTGAT 813  
Qy 538 TCAGACCGGTGGGGCTTCCGGCTGCACAGTGTCTGACAGAAATTCAGAAATTGGG 597  
Db 814 TCAGACCGGTGGGGCTTCCGGCTGCACAGTGTCTGACAGAAATTCAGAAATTGGG 873  
Qy 598 AAATGAGCTACAGCAGAAAGATGCTGCTGCTGCGGACCAAGTGTGACTTACACAGGA 657  
Db 874 AAATGAGCTACAGCAGAAAGATGCTGCTGCTGCGGACCAAGTGTGACTTACACAGGA 933  
Qy 658 AATGTTCTCTACGTCTTCTCTCTCCCTGTAAGAGATGACAGCGAAAGAGAGAGCT 717  
Db 934 AATGTTCTCTACGTCTTCTCTCTCCCTGTAAGAGATGACAGCGAAAGAGAGAGCT 993  
Qy 718 GTACAGAGATTATCTCCAGGAGTGGCCAGACTCCAGGCTCCGGTACAGAGATCATG 777  
Db 994 GTACAGAGATTATCTCCAGGAGTGGCCAGACTCCAGGCTCCGGTACAGAGATCATG 1053  
Qy 778 TCCGACCAAGAAAGCTGTTTTCATCATTTGACGGTTTGATGACCTGCGCTCTGCTC 837  
Db 1054 TCCGACCAAGAAAGCTGTTTTCATCATTTGACGGTTTGATGACCTGCGCTCTGCTC 1113  
Qy 838 AACATGACACAAAGCTCTGCAAGAGCTGAGAGACCTTCCTTACCTCAT 897  
Db 1114 AACATGACACAAAGCTCTGCAAGAGCTGAGAGACCTTCCTTACCTCAT 1173  
Qy 898 CGCAGTCTGCTGAGGAAAGTCTGCTCCCTGAGTCCCTGATGATGACCGTACAGAGC 957  
Db 1174 CGCAGTCTGCTGAGGAAAGTCTGCTCCCTGAGTCCCTGATGATGACCGTACAGAGC 1233  
Qy 958 GTGGCACAAGAGACTCAAGTCAAGAGTCTGCTCCCTTACCTGTTAGTAGAGA 1017  
Db 1234 GTGGCACAAGAGACTCAAGTCAAGAGTCTGCTCCCTTACCTGTTAGTAGAGA 1293  
Qy 1018 ATCTCCGGGAAACAAAGATCCATTGCTCTTGAAGCCGGGATTGTGAGCATCAAG 1077  
Db 1294 ATCTCCGGGAAACAAAGATCCATTGCTCTTGAAGCCGGGATTGTGAGCATCAAG 1353  
Qy 1078 ACACAAAGGTTGCGGCGATCAATGAACAACGCTGAGCTGACCAAGTCCAGGTGCGC 1137  
Db 1354 ACACAAAGGTTGCGGCGATCAATGAACAACGCTGAGCTGACCAAGTCCAGGTGCGC 1413  
Qy 1138 GCCGTGAGCTCTCATCTGCTGCGCTGACAGTCAAGAGCTGAGGAGAGAGCTC 1197  
Db 1414 GCCGTGAGCTCTCATCTGCTGCGCTGACAGTCAAGAGCTGAGGAGAGAGCTC 1473  
Qy 1198 GCCCCTTTCACAAACGCTCAAGGCTGACAGCCGCTTTTGTGTTTTCATCAAGTCA 1257  
Db 1474 GCCCCTTTCACAAACGCTCAAGGCTGACAGCCGCTTTTGTGTTTTCATCAAGTCA 1533  
Qy 1258 CCTGAGGCGTGTCCGGGCTGCTCAATCTGAGGAAAGGTTGCTGTAAGGCTTC 1317  
Db 1534 CCTGAGGCGTGTCCGGGCTGCTCAATCTGAGGAAAGGTTGCTGTAAGGCTTC 1593  
Qy 1318 TGCCGTATGCTGTGAGAGAGTGTGAATAGAAAGTCAAGTGTGATGTGACAGCTC 1377  
Db 1594 TGCCGTATGCTGTGAGAGAGTGTGAATAGAAAGTCAAGTGTGATGTGACAGCTC 1653

Qy 1378 AATGTTCAAGAGCTGCGGGAAGTCTGAGCTCCGTGCTGTTTCAATGAACATCTTCTC 1437  
Db 1654 AATGTTCAAGAGCTGCGGGAAGTCTGAGCTCCGTGCTGTTTCAATGAACATCTTCTC 1713  
Qy 1438 CCAGACAGCACTGTGAGAGTACTACACTTCTTCCACTCATGCTCCAGAGCTTCTGT 1497  
Db 1714 CCAGACAGCACTGTGAGAGTACTACACTTCTTCCACTCATGCTCCAGAGCTTCTGT 1773  
Qy 1498 GCCGCTTGTACTACGTGTGAGAGGCTTGAATTCAGACAGCTCTGCTCTGTAC 1557  
Db 1774 GCCGCTTGTACTACGTGTGAGAGGCTTGAATTCAGACAGCTCTGCTCTGTAC 1833  
Qy 1558 GTTGAAGAACAAAGAGTCCATGAGAGCTTAAACAGAGAGCTTCCATATCACTGCTT 1617  
Db 1834 GTTGAAGAACAAAGAGTCCATGAGAGCTTAAACAGAGAGCTTCCATATCACTGCTT 1893  
Qy 1618 TGATGAAGAGCTTCTTGTGAGCTCTGAGACGAAGTGAAGAGGCACTGAGAGTTC 1677  
Db 1894 TGATGAAGAGCTTCTTGTGAGCTCTGAGACGAAGTGAAGAGGCACTGAGAGTTC 1953  
Qy 1678 CTGCTGGCTGTCCCGTTCCTCGGGGTGAAGCAAGAGCTTTCAGTGGGTCTCTG 1737  
Db 1954 CTGCTGGCTGTCCCGTTCCTCGGGGTGAAGCAAGAGCTTTCAGTGGGTCTCTG 2013  
Qy 1738 TTGGGTACAGAGCTTAATGCCACCACTCCAGAGACACCTCTGAGAGCTTCACTGTCTT 1797  
Db 2014 TTGGGTACAGAGCTTAATGCCACCACTCCAGAGACACCTCTGAGAGCTTCACTGTCTT 2073  
Qy 1798 TTGAGACTCAAGAACAAAGAGTGTGCTGCTGAGCTTAAACAGCTTCAAGAAAGTGG 1857  
Db 2074 TTGAGACTCAAGAACAAAGAGTGTGCTGCTGAGCTTAAACAGCTTCAAGAAAGTGG 2133  
Qy 1858 CTTCGATTAACAGAGCTGAGCTTGAATGATCTTCTTCTGCTCAGACCTGTCCG 1917  
Db 2134 CTTCGATTAACAGAGCTGAGCTTGAATGATCTTCTTCTGCTCAGACCTGTCCG 2193  
Qy 1918 TATTTGCGGAAATTCGGGTGAGATGTCAAAAGGATCTTCCCAAGATGAGTCCGCTAG 1977  
Db 2194 TATTTGCGGAAATTCGGGTGAGATGTCAAAAGGATCTTCCCAAGATGAGTCCGCTAG 2253  
Qy 1978 GCATGCTCTGTGTCCTCTATGATGATGAGGAGTAAAGACCTCATTTAGAGAGAGGGAA 2037  
Db 2254 GCATGCTCTGTGTCCTCTATGATGATGAGGAGTAAAGACCTCATTTAGAGAGAGGGAA 2313  
Qy 2038 GATTTCTGCTCATGCTTGGACCCACCACTGCGGAGCTGAGCTTGGGAGCAGC 2097  
Db 2314 GATTTCTGCTCATGCTTGGACCCACCACTGCGGAGCTGAGCTTGGGAGCAGC 2373  
Qy 2098 ATCTGACAGAGCGGGCCATGAAGACCTGTGTGTCMAAGCTGAGGATCCACCTGCAAG 2157  
Db 2374 ATCTGACAGAGCGGGCCATGAAGACCTGTGTGTCMAAGCTGAGGATCCACCTGCAAG 2433  
Qy 2158 ATACAGACCTGATGTTAGAAATGACACAGATTACCTCGGTGTGACCACTCTGAGGA 2217  
Db 2434 ATACAGACCTGATGTTAGAAATGACACAGATTACCTCGGTGTGACCACTCTGAGGA 2493  
Qy 2218 ATCTCATGAGCAACCGTAACTTAAGATCCCTCAACTTGGAGGACCACTGAAGAA 2277  
Db 2494 ATCTCATGAGCAACCGTAACTTAAGATCCCTCAACTTGGAGGACCACTGAAGAA 2553  
Qy 2278 GAGGATGAAGAGTGGGTGTGAAGCTTAAACCAACCAAAATGTTGTGAGTCTTGG 2337  
Db 2554 GAGGATGAAGAGTGGGTGTGAAGCTTAAACCAACCAAAATGTTGTGAGTCTTGG 2613  
Qy 2338 AGGCTGATGCTGTGATGTAACCATGCTGTACCTGAAGATCTCCAAATCTTAAG 2397  
Db 2614 AGGCTGATGCTGTGATGTAACCATGCTGTACCTGAAGATCTCCAAATCTTAAG 2673  
Qy 2398 ACTTCCCAAGCTGAAGATCTGTAGGCTGACAGAAACAAAGTGAAGACCAAGAGATA 2457  
Db 2674 ACTTCCCAAGCTGAAGATCTGTAGGCTGACAGAAACAAAGTGAAGACCAAGAGATA 2733



QY 2458 ATGCTCTAGTATGCTTGGAGAGTCTCCAGTGCCTGACAGAGCTGATCTGAG 2517  
 DB 2734 AGCGCTCTAGTATGCTTGGAGAGTCTCCAGTGCCTGACAGAGCTGATCTGAG 2793  
 QY 2518 GACTGTGGATCAACAGCCAGGGTTGCCAGAGTCTGGCCTCAGCCCTGTCAGCAACGG 2577  
 DB 2794 GACTGTGGATCAACAGCCAGGGTTGCCAGAGTCTGGCCTCAGCCCTGTCAGCAACGG 2853  
 QY 2578 AGCTTGACACACCTGTGGCTATCCAAACAAGCCGGGGAAAGAGGTAAATCTACTG 2637  
 DB 2854 AGCTTGACACACCTGTGGCTATCCAAACAAGCCGGGGAAAGAGGTAAATCTACTG 2913  
 QY 2638 TGTGATCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATCTGATCATGTCAC 2697  
 DB 2914 TGTGATCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATCTGATCATGTCAC 2973  
 QY 2698 CTGACACGGCTGGCTGTGGTTTCTTGCACTGGCCTTAAAGGTAATCATGGCTGACG 2757  
 DB 2974 CTGACACGGCTGGCTGTGGTTTCTTGCACTGGCCTTAAAGGTAATCATGGCTGACG 3033  
 QY 2758 CACCTGAGCCTTAGATGAACCTGTGGAAAGCAATGGCGTGAAGCTTCTGTGGAGGTC 2817  
 DB 3034 CACCTGAGCCTTAGATGAACCTGTGGAAAGCAATGGCGTGAAGCTTCTGTGGAGGTC 3093  
 QY 2818 ATGAGAGAACCATCTTGTCACTCCAGGACCTGAGTTGTAAAGTGTCACTCACCGCC 2877  
 DB 3094 ATGAGAGAACCATCTTGTCACTCCAGGACCTGAGTTGTAAAGTGTCACTCACCGCC 3153  
 QY 2878 GGTGTCTGTGAGAGTCTGTCTGTGTGATCTTGAGAGACAGACACCTGAAGAGCTGAT 2937  
 DB 3154 GGTGTCTGTGAGAGTCTGTCTGTGTGATCTTGAGAGACAGACACCTGAAGAGCTGAT 3213  
 QY 2938 CTCAGGACCAATGCGCTGGGTGAAGGTGGGGTGTGGTGGCTGTGAGAGGACTGAAGCA 2997  
 DB 3214 CTCAGGACCAATGCGCTGGGTGAAGGTGGGGTGTGGTGGCTGTGAGAGGACTGAAGCA 3273  
 QY 2998 AAGAACAGTGTCTGACGAGACTCGGGTGTGAAGGAGTGTGACTGATCTGATTTGCTGT 3057  
 DB 3274 AAGAACAGTGTCTGACGAGACTCGGGTGTGAAGGAGTGTGACTGATCTGATTTGCTGT 3333  
 QY 3058 GAGGACACTCTCTTGGCCCTTCTGTCGAAACCGGCACTTGACCAAGTCTAAACCTGGTGCAG 3117  
 DB 3334 GAGGACACTCTCTTGGCCCTTCTGTCGAAACCGGCACTTGACCAAGTCTAAACCTGGTGCAG 3393  
 QY 3118 AATACTTCAGTCCCAAGGAATGATGAAGCTGTGGGCTTGGCCCTTCCCAAGCTCT 3177  
 DB 3394 AATACTTCAGTCCCAAGGAATGATGAAGCTGTGGGCTTGGCCCTTCCCAAGCTCT 3453  
 QY 3178 AACTTACAGATTAATGGCTGTGGAATGGCAGTACCTGTGCAATAAAGGAAGCTGCTG 3237  
 DB 3454 AACTTACAGATTAATGGCTGTGGAATGGCAGTACCTGTGCAATAAAGGAAGCTGCTG 3513  
 QY 3238 GAGGAAGTGCAGCTACTCAAGCCCGGAGTCTGTAATGACGGTAGTGGCAATCTTTTAT 3297  
 DB 3514 GAGGAAGTGCAGCTACTCAAGCCCGGAGTCTGTAATGACGGTAGTGGCAATCTTTTAT 3573  
 QY 3298 GAAGATGACCGACAC 3312  
 DB 3574 GAAGATGACCGGTAC 3588

RESULT 5  
 ADA45220  
 ID ADA45220 standard; DNA; 3830 BP.

AC ADA45220;  
 XX 20-NOV-2003 (first entry)  
 DT  
 XX Human MATER splice variant DNA SEQ ID 3.  
 DE human; MATER; maternal antigen that embryos require; NTP-ase;  
 KW nucleoside triphosphate-ase; apoptosis; antiinfertility; contraceptive;

KW antiinflammatory; immunosuppressive; gynaecological; endometriosis;  
 KW ovarian dysfunction; autoimmune premature ovarian failure;  
 KW autoimmune disease; contraceptive; autoantibody; ds; gene.  
 OS Homo sapiens.  
 PH Key Location/Qualifiers  
 FT CDS 1..3432  
 FT /\*tag= a  
 FT /product= "splice variant of MATER"  
 PN EP1285964-A2.  
 PD 26-FEB-2003.  
 PD 12-JUL-2002; 2002EP-00090246.  
 PR 10-AUG-2001; 2001DE-01039874.  
 PA (SCHD ) SCHERING AG.  
 PI Weiss B, Lessl M, Peters-Kottig M, Beckmann G;  
 DR WP1; 2003-302814/30.  
 DR P-PSDB; ADA45221.  
 PT New nucleic acid for human maternal antigen that embryos require protein,  
 useful e.g. in diagnosis and treatment of female infertility.  
 PS Claim 1; Page 17-18; 31pp; German.  
 CC This invention describes a novel human MATER (maternal antigen that  
 CC embryos require) protein. MATER is an NTP(nucleoside triphosphate)-ase  
 CC associated with apoptosis, a defect in MATER activity causes growth  
 CC arrest at the 2-cell stage of fertilized eggs. mRNA transcribed from the  
 CC MATER gene is found predominantly in ovary, testis and placenta. The  
 CC products of the invention have antiinfertility, contraceptive,  
 CC antiinflammatory, immunosuppressive and gynaecological activity.  
 CC Effectors of the MATER polypeptide are used to treat infertility.  
 CC associated with endometriosis and also ovarian dysfunction, autoimmune  
 CC premature ovarian failure, inflammation, autoimmune diseases and female  
 CC infertility, and as contraceptives. Measurements of autoantibodies  
 CC against MATER in body samples of MATER protein or mRNA in eggs is useful  
 CC for diagnosis of female infertility, particularly by detecting mutations  
 CC in the MATER gene with a DNA chip. This sequence represents a splice  
 CC variant of human MATER in which exon 4 of the usual form is absent.  
 SQ Sequence 3830 BP; 943 A; 946 C; 1041 G; 900 T; 0 U; 0 Other;

Query Match 77.8%; Score 3139; DB 8; Length 3830;  
 Best Local Similarity 96.0%; Pred. No. 0;  
 Matches 3282; Conservative 0; Mismatches 30; Indels 105; Gaps 2;

QY 1 ATGGAAGAGACAATGCTTACCTTTCCAGTACGAGGCTGCAATGTCTCTATGAG 60  
 DB 1 ATGGAAGAGACAATGCTTACCTTTCCAGTACGAGGCTGCAATGTCTCTATGAG 60  
 QY 61 CTAGACAAGGAAGATTTCAGACATTCAGGAATTACTTAAAGAAATCTTCGAATCG 120  
 DB 61 CTAGACAAGGAAGATTTCAGACATTCAGGAATTACTTAAAGAAATCTTCGAATCG 120  
 QY 121 ACCAGATGCTATTTCCAGATTTGAAGTGAAGATGCCAAGCGAATGCTGGCACTC 180  
 DB 121 ACCAGATGCTATTTCCAGATTTGAAGTGAAGATGCCAAGCGAATGCTGGCACTC 180  
 QY 181 CTTTGATGAGTATTAATGAGCATGCTGGCTGAGCTAGTCCATTAGCATCTTTGAA 240  
 DB 181 CTTTGATGAGTATTAATGAGCATGCTGGCTGAGCTAGTCCATTAGCATCTTTGAA 240  
 QY 241 AACATGAACCTGGCAACCTCTCGAGAAAGCGAGATGATGAAAAAATTTACAA 300  
 DB 241 AACATGAACCTGGCAACCTCTCGAGAAAGCGAGATGATGAAAAAATTTACAA 300



Db 2461 TTGACCCATGCTGTTACCTGTAAGATCTCCAAATCTTACGACCTCCCGAGCCCTGAAA 2520  
Qy 2416 TCTCTGAGCCTGGCAGAAACAGGTGACAGACCGAGGAGTAATGCTCTAGTATGATGCC 2475  
Db 2521 TCTCTGAGCCTGGCAGAAACAGGTGACAGACCGAGGAGTAATGCTCTAGTATGATGCC 2580  
Qy 2476 TTGAGAGTCTCCAGTGGCCCTGAGAGAGTGTATCTGAGAGACTGTGGATCAACAGCC 2535  
Db 2581 TTGAGAGTCTCCAGTGGCCCTGAGAGAGTGTATCTGAGAGACTGTGGATCAACAGCC 2640  
Qy 2536 ACGGGTTGCCAGAGTCTGGCTCAGCCCTCTGAGCAACCGAGACTTGAACACCTGTGC 2595  
Db 2641 ACGGGTTGCCAGAGTCTGGCTCAGCCCTCTGAGCAACCGAGACTTGAACACCTGTGC 2700  
Qy 2596 CTATCCAAACAACAGCTGGGGAGAGAGTGTAAATCTACTGTGTGATTCATGAGAGCTT 2655  
Db 2701 CTATCCAAACAACAGCTGGGGAGAGAGTGTAAATCTACTGTGTGATTCATGAGAGCTT 2760  
Qy 2656 CCCCACTGAGTCTGAGAGAGCTGATGCTGAATCAGTCCACCTGAGACAGCGCTGGCTGT 2715  
Db 2761 CCCCACTGAGTCTGAGAGAGCTGATGCTGAATCAGTCCACCTGAGACAGCGCTGGCTGT 2820  
Qy 2716 GGTTCCTTGTGACCTTGGCTTAATGGGTAACTGATGCTGAGAGCAGCTGAGCTTACATG 2775  
Db 2821 GGTTCCTTGTGACCTTGGCTTAATGGGTAACTGATGCTGAGAGCAGCTGAGCTTACATG 2880  
Qy 2776 AACCTGTGGAAGAACAATAGGGGTGAAGCTTCTGTGCGAGGTGATGAGAGAAACAATCTTGT 2835  
Db 2881 AACCTGTGGAAGAACAATAGGGGTGAAGCTTCTGTGCGAGGTGATGAGAGAAACAATCTTGT 2940  
Qy 2836 CATCTCCAGAGACTGAGAGTGTGTAAGTGTCACTCAACCGCGCGCTGTGTGAGAGTCTG 2895  
Db 2941 CATCTCCAGAGACTGAGAGTGTGTAAGTGTCACTCAACCGCGCGCTGTGTGAGAGTCTG 3000  
Qy 2896 TCTCTGTGATCTGAGAGACAGACACTGAAAGAGCTGTGATCTGACGAACTAATGCCCTTG 2955  
Db 3001 TCTCTGTGATCTGAGAGACAGACACTGAAAGAGCTGTGATCTGACGAACTAATGCCCTTG 3060  
Qy 2956 GGTGACGGTGGGGTGTGCTGCGCTGTGCGAGGAGCTGAAAGAAACAGTGTCTGACG 3015  
Db 3061 GGTGACGGTGGGGTGTGCTGCGCTGTGCGAGGAGCTGAAAGAAACAGTGTCTGACG 3120  
Qy 3016 AGACTCGGGTGAAGAGCATGTGACTGACTTCTGATTCCTGTGAGGCACTCTCTTGGCC 3075  
Db 3121 AGACTCGGGTGAAGAGCATGTGACTGACTTCTGATTCCTGTGAGGCACTCTCTTGGCC 3180  
Qy 3076 CTTTCTGTGAACCGGCATCTGACAGTCTAAACCTGTGCGAGATTAATTCACTGCCAAA 3135  
Db 3181 CTTTCTGTGAACCGGCATCTGACAGTCTAAACCTGTGCGAGATTAATTCACTGCCAAA 3240  
Qy 3136 GGAATGATGAAGCTGTGTTGGCCCTTTCGCTGCTCCAGCTTAACCTTACAGATAATTGGG 3195  
Db 3241 GGAATGATGAAGCTGTGTTGGCCCTTTCGCTGCTCCAGCTTAACTTACAGATAATTGGG 3300  
Qy 3196 CTGTGGAATGAGCAGTACCTCTGTGCAAAATGAAGAGCTGTGAGAGAGTCAAGTACTC 3255  
Db 3301 CTGTGGAATGAGCAGTACCTCTGTGCAAAATGAAGAGCTGTGAGAGAGTCAAGTACTC 3360  
Qy 3256 AAGCCCCGAGTCTGTAATGAGAGGTGTTGGCATTTCTTTGATGAAGATGACCGGATC 3312  
Db 3361 AAGCCCCGAGTCTGTAATGAGAGGTGTTGGCATTTCTTTGATGAAGATGACCGGATC 3417

RESULT 6  
AAD41224  
ID AAD41224 standard; cDNA, 3489 BP.

XX AAD41224;  
XX AC  
XX 30-OCT-2002 (first entry)  
XX DT  
XX Human EMBRY-1 cDNA.  
XX DE  
XX

KW Human; embryogenesis associated protein; AIDS; reproductive disorder;  
KW infertility; endometriosis; endometrial tumour; inflammatory disorder;  
KW autoimmune disorder; acquired immune deficiency syndrome; transgenic;  
KW ovarian tumour; contact dermatitis; placenta disorder; preeclampsia;  
KW EMBRY-1; allergy; gene therapy; gene; ss.  
XX Homo sapiens.  
OS  
XX  
FH Key 1..3489  
FT CDS /tag= a  
FT /product= "EMBRY-1 protein"  
PN  
XX  
PD 20-JUN-2002.  
XX  
XX 14-NOV-2001; 2001MO-US043956.  
PF  
XX 15-NOV-2000; 2000US-0249407P.  
PR  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
XX Rankumar J, Arvizu C;  
PI  
XX  
XX MPI; 2002-537629/57.  
DR P-PSDB; AAE25053.  
PT New polypeptides of human embryogenesis associated proteins for screening  
PT modulators useful for treating or preventing disorders e.g.  
PT endometriosis, infertility, allergy, preeclampsia.  
XX  
XX  
PS Claim 58; Page 95-96; 97pp; English.

CC The invention relates to human embryogenesis associated proteins (EMBRY)  
CC and nucleic acid molecules encoding such proteins. EMBRY sequences are  
CC useful for screening modulators useful for treating or preventing  
CC disorders associated with abnormal expression of EMBRY. The disorders  
CC treated include reproductive disorders such as infertility,  
CC endometriosis, endometrial or ovarian tumour, autoimmune/inflammatory  
CC disorder such as acquired immune deficiency syndrome (AIDS), allergies,  
CC contact dermatitis; disorders of the placenta such as preeclampsia,  
CC abruptio placentae etc. Sequences of the invention are also useful for  
CC analysing a proteome of a tissue or a cell type. EMBRY proteins are  
CC useful as immunogens for preparing antibodies. Polynucleotides of the  
CC invention are useful for creating knockin humanised animals or transgenic  
CC animals to model human diseases. They are also used in gene therapy. The  
CC present sequence is human EMBRY-1 cDNA  
XX  
XX

Sequence 3489 BP; 854 A; 881 C; 961 G; 793 T; 0 U; 0 Other;

Query March 77.7%; Score 3133.6; DB 6; Length 3489;  
Best Local Similarity 95.2%; Pred. No. 0;  
Matches 3308; Conservative 0; Mismatches 4; Indels 162; Gaps 1;

Qy 1 ATGGAAGAGCAAAATGCTACCTTTCCAGTACCGGGCTGCAATGGTCTATGAG 60  
Db 1 ATGGAAGAGCAAAATGCTACCTTTCCAGTACCGGGCTGCAATGGTCTATGAG 60  
Qy 61 CTGACCAAGGAAGATTTCAGACATTCAGGAATTACTTAAGAAATCTTCGAATCG 120  
Db 61 CTGACCAAGGAAGATTTCAGACATTCAGGAATTACTTAAGAAATCTTCGAATCG 120  
Qy 121 ACCGATGCTCTATTTCCAGAGTTTGAATGAGAAATCCAAACGTGAATGCTGGCACTC 180  
Db 121 ACCGATGCTCTATTTCCAGAGTTTGAATGAGAAATCCAAACGTGAATGCTGGCACTC 180  
Qy 181 CTTTGCATGAGATTTATGAGAGCATCGCTGGCTGAGCTAGTCCATTAGCATCTTGA 240  
Db 181 CTTTGCATGAGATTTATGAGAGCATCGCTGGCTGAGCTAGTCCATTAGCATCTTGA 240  
Qy 241 AACATGAACCTGCGAACCTCTCGGAGAGGCAAGGAGTACATGAATA----- 289

241 AACATGAACCTGCAACCTCTCGAGAGAGCAGGGATGATGATAAAAAATTCACCAGAA 300  
290 ----- 289  
301 GATCCTGAAGACGATGATGACCAAGGACCAAGCAAGAAAAAGTGCAGAAAAATAAA 360  
290 ----- 289  
361 TATGGCATGACTAAGCTTATCTTGGGGGTGTCTGACATCTCTGACTCGAATAATAACAC 420  
290 ----- AAATTTCAAGGCTATGAGCAAGAGGT 318  
421 AAGTATGTTGAATTCATTCTTTTGAGAAATTTCAAGGCTATGAGCAAGAGGT 480  
319 GCCACAGCAGCAGAGCAGAGAAACAGAAATTTTCAAGGCTATGAGCAAGAGGTGCC 378  
481 GCCACAGCAGCAGAGCAGAGAAACAGAAATTTTCAAGGCTATGAGCAAGAGGTGCC 540  
379 ACAGCAGCAGAGCAGAGAAACAGAGCAATGAGGTCACATGAGGACTCAAGAGTAC 438  
541 ACAGCAGCAGAGCAGAGAAACAGAGCAATGAGGTCACATGAGGACTCAAGAGTAC 600  
439 GTGATGACCAAAATTCGCTGAGAGAGAGATGATACGTAGTATTTTGAACACCTGCTGT 498  
601 GTGATGACCAAAATTCGCTGAGAGAGAGATGATACGTAGTATTTTGAACACCTGCTGT 660  
499 GACTGGCCGGAATGCAAACTTTGGCTGTCTTTTGAATTCAGACCGGTGGGGCTTCCGG 558  
661 GACTGGCCGGAATGCAAACTTTGGCTGTCTTTTGAATTCAGACCGGTGGGGCTTCCGG 720  
559 CCTGCGACGCTGCTTCTGACGAGAAAGTCAGGAATTCGAGGCTTACCCAGAAAG 618  
721 CCTGCGACGCTGCTTCTGACGAGAAAGTCAGGAATTCGAGGCTTACCCAGAAAG 780  
619 ATCGTCTGTGCTGGGCGCAGAGTGAATCTTACAGGGAATTTCTCTACCTCTTCTTC 678  
781 ATCGTCTGTGCTGGGCGCAGAGTGAATCTTACAGGGAATTTCTCTACCTCTTCTTC 840  
679 CTCCCCGTTAGAGAGATGACGCGAAGAAAGAGAGAGATGTCACAGAGTTCACTCCAG 738  
841 CTCCCCGTTAGAGAGATGACGCGAAGAAAGAGAGAGATGTCACAGAGTTCACTCCAG 900  
739 GAGTGGCCAGATCCCAAGCTCCGAGTACGAGATCATGTCGCGACCAAGAAAGGCTTGG 798  
901 GAGTGGCCAGATCCCAAGCTCCGAGTACGAGATCATGTCGCGACCAAGAAAGGCTTGG 960  
799 TTCAATCATTTGACGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 858  
961 TTCAATCATTTGACGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
859 AAAAGCTGGGCTGAGAGAGAGCTCCGTTACCTCATACGAGTCTGCTGAGAGAGTGC 918  
1021 AAAAGCTGGGCTGAGAGAGAGCTCCGTTACCTCATACGAGTCTGCTGAGAGAGTGC 1080  
919 CTGCTCCCTGATGATCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 978  
1081 CTGCTCCCTGATGATCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
979 TCAGAGTCTGTGCTCCCGTTCATGTTAGTGAAGAAATCTCCGGGGAAACAAAGAAATC 1038  
1141 TCAGAGTCTGTGCTCCCGTTCATGTTAGTGAAGAAATCTCCGGGGAAACAAAGAAATC 1200  
1039 CACTTGTCTTGAAGCGGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1098  
1201 CACTTGTCTTGAAGCGGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
1099 ATGAACAACCGTGAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1158  
1261 ATGAACAACCGTGAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
1159 GTGGCCCTGACAGCTGACAGAGCTGTGTGGGGAGAGAGTGTGCTTCAACCAAGCGCTC 1218  
1321 GTGGCCCTGACAGCTGACAGAGCTGTGTGGGGAGAGAGTGTGCTTCAACCAAGCGCTC 1380

1219 AAGGCTGACAGCGCCCTTTTGTGTTTCATGAGCTCAACCCCTGAGAGAGTGTCCGGGC 1278  
1381 AAGGCTGACAGCGCCCTTTTGTGTTTCATGAGCTCAACCCCTGAGAGAGTGTCCGGGC 1440  
1279 TGTCTCAATCTGAGAGAAAGTGTCTTGAAGAGCTTCTGCGTATGCTGTGAGAGGA 1338  
1441 TGTCTCAATCTGAGAGAAAGTGTCTTGAAGAGCTTCTGCGTATGCTGTGAGAGGA 1500  
1339 GTGTGAATGAGAGTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1398  
1501 GTGTGAATGAGAGTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
1339 TGTGAGCTCCGCTGTCTGTTTCAATGAACATCCCTTCCAGACAGCAGCAGTGTGAGAG 1458  
1561 TGTGAGCTCCGCTGTCTGTTTCAATGAACATCCCTTCCAGACAGCAGCAGTGTGAGAG 1620  
1459 TACTACACCTTCTTCACTCAGTCTCAGAGACTTCTGCGCGCTTGTACTACGTGTTA 1518  
1621 TACTACACCTTCTTCACTCAGTCTCAGAGACTTCTGCGCGCTTGTACTACGTGTTA 1680  
1519 GAGGCTGGAATGAGAGCAGCTTCTGCGCTCTGTAAGTGAAGACAAAGAGTCC 1578  
1681 GAGGCTGGAATGAGAGCAGCTTCTGCGCTCTGTAAGTGAAGACAAAGAGTCC 1740  
1579 ATGAGCTTAAACAGGAGGCTTCAATATCAATCGCTTTGATGATGATGATGATGATGAT 1638  
1741 ATGAGCTTAAACAGGAGGCTTCAATATCAATCGCTTTGATGATGATGATGATGATGAT 1800  
1639 GAGCTGTGAGAGAGAGTGAAGAGGAGCAGCTGAGAGTCTGCTGAGGCTGTCCGGTCC 1698  
1801 GAGCTGTGAGAGAGAGTGAAGAGGAGCAGCTGAGAGTCTGCTGAGGCTGTCCGGTCC 1860  
1699 CTGGGGGTGAAGCAGAACTTCTGACATGAGTCTCTGTTGGGTCAAGACCTAATGCC 1758  
1861 CTGGGGGTGAAGCAGAACTTCTGACATGAGTCTCTGTTGGGTCAAGACCTAATGCC 1920  
1759 ACCACCCAGAGAGACCTCTGAGAGGCTTCCATGCTTTTTCAGAGCTCAAGACAAAG 1818  
1921 ACCACCCAGAGAGACCTCTGAGAGGCTTCCATGCTTTTTCAGAGCTCAAGACAAAG 1980  
1819 TTTGTTGCTTGGGATTTAAACAGCTTCCAAAGAGTGTGCTTCCGATTAACAGAACTTG 1878  
1981 TTTGTTGCTTGGGATTTAAACAGCTTCCAAAGAGTGTGCTTCCGATTAACAGAACTTG 2040  
1879 GACTTGAATGATCTTCTTCTGCTCCAGACATGCTCCGATTTTTCGGAATAATTCGGGTG 1938  
2041 GACTTGAATGATCTTCTTCTGCTCCAGACATGCTCCGATTTTTCGGAATAATTCGGGTG 2100  
1939 GATGTCAAAAGGATCTTCCAAAGAGTGTCCGCTGAGGAGTCTGTGCTCCCTTA 1998  
2101 GATGTCAAAAGGATCTTCCAAAGAGTGTCCGCTGAGGAGTCTGTGCTCCCTTA 2160  
1999 TGAATGGGGATTAAGACCTCTATGAGAGAGAGTGTGGAAGATTTCTGCTCCATGCTTGGC 2058  
2161 TGAATGGGGATTAAGACCTCTATGAGAGAGAGTGTGGAAGATTTCTGCTCCATGCTTGGC 2220  
2059 ACCACCCAGACCTGCGGAGCTGAGACCTGCGGAGAGCATCTGACAGAGCGGGCCATG 2118  
2221 ACCACCCAGACCTGCGGAGCTGAGACCTGCGGAGAGCATCTGACAGAGCGGGCCATG 2280  
2119 AAGACCTGTGTGCCAAGTGAAGCATCCACTGCAAGATACAGACCTGATGTTTGA 2178  
2281 AAGACCTGTGTGCCAAGTGAAGCATCCACTGCAAGATACAGACCTGATGTTTGA 2340  
2179 AATGACAGATTACCCCTGTGTGACAGACCTCTGGAAGATCCGATAGGCCAACCTGTAC 2238  
2341 AATGACAGATTACCCCTGTGTGACAGACCTCTGGAAGATCCGATAGGCCAACCTGTAC 2400  
2239 CTAAGATCCCTCAACTTGGAGAGCAGCAGCAGCTGGAAGAGAGATGTAAGATGAGCTGT 2238  
2401 CTAAGATCCCTCAACTTGGAGAGCAGCAGCAGCTGGAAGAGAGATGTAAGATGAGCTGT 2460

```

QY 2299 GAAGCCTTAAAAACCCAAAAATGTTGTGAGTCTTTGAGGCTGGAATGCTGATG 2358
DB 2461 GAAGCTTAAAAACCCAAAAATGTTGTGAGTCTTTGAGGCTGGAATGCTGATG 2520
QY 2359 ACCCATGCTTGTACTGAAAGTCTCCCAATCCTTACGACCTCCCGACGCTGAATCT 2418
DB 2521 ACCCATGCTTGTACTGAAAGTCTCCCAATCCTTACGACCTCCCGACGCTGAATCT 2580
QY 2419 CTGAGCCTGGAGAAAAGAGTGAAGAGCCAGAGGAGTAAATGCTCTGAGTATGCTTG 2478
DB 2561 CTGAGCCTGGAGAAAAGAGTGAAGAGCCAGAGGAGTAAATGCTCTGAGTATGCTTG 2640
QY 2479 AGAGTCTCCAGATGCGCTCTGAGAGCTGATATCTGAGAGCTGTGAGTCAAGCCACG 2538
DB 2641 AGAGTCTCCAGATGCGCTCTGAGAGCTGATATCTGAGAGCTGTGAGTCAAGCCACG 2700
QY 2539 GGTTCGCAAGTCTGTGCTCAGCCTCTGTCAGCAACCGAGCTTGAACAACCTGTGCTTA 2598
DB 2701 GGTTCGCAAGTCTGTGCTCAGCCTCTGTCAGCAACCGAGCTTGAACAACCTGTGCTTA 2760
QY 2599 TCCAAACAAGCCTGGGAGAGAGGTTAAATCTACTGTGTGATGATGAGGCTTCC 2658
DB 2761 TCCAAACAAGCCTGGGAGAGAGGTTAAATCTACTGTGTGATGATGAGGCTTCC 2820
QY 2659 CACTGTAGTCTGCAAGGCTGATGCTGATCAAGTCCACTGCAACCGCTGCTGTGTGT 2718
DB 2821 CACTGTAGTCTGCAAGGCTGATGCTGATCAAGTCCACTGCAACCGCTGCTGTGTGT 2880
QY 2719 TTTCTTGACATTTGCTGCTTATGGGTAACTCATGCTGACGACCTGAGCTTACATGAA 2778
DB 2881 TTTCTTGACATTTGCTGCTTATGGGTAACTCATGCTGACGACCTGAGCTTACATGAA 2940
QY 2779 CCTGTGAAAGCAATAGGCGTGAAGCTTCTGTGCGAGGTCAAGAGAAACAATCTTGTAT 2838
DB 2941 CCTGTGAAAGCAATAGGCGTGAAGCTTCTGTGCGAGGTCAAGAGAAACAATCTTGTAT 3000
QY 2839 CTCCAGAACCTGAGATTGTAAAGTGTATCTCAACCGCGCTGTGTGAAGTCTGCC 2898
DB 3001 CTCCAGAACCTGAGATTGTAAAGTGTATCTCAACCGCGCTGTGTGAAGTCTGCC 3060
QY 2899 TGTGTGATCTGAGAGAGCAGACACTTGAAGAGCTGTGATCTCAAGGAAACAATGCTTGGGT 2958
DB 3061 TGTGTGATCTGAGAGAGCAGACACTTGAAGAGCTGTGATCTCAAGGAAACAATGCTTGGGT 3120
QY 2959 GACGCTGGAGGTTGCTGCGCTGTGCGAGGAGCTGAAGAAAAGACAGTCTCTGACGAGA 3018
DB 3121 GACGCTGGAGGTTGCTGCGCTGTGCGAGGAGCTGAAGAAAAGACAGTCTCTGACGAGA 3180
QY 3019 CTCGGGTTGAAGGACATGTGAAGCTTCTGATGCTGTGAGGCACTCTCCCTTGGCCCTT 3078
DB 3181 CTCGGGTTGAAGGACATGTGAAGCTTCTGATGCTGTGAGGCACTCTCCCTTGGCCCTT 3240
QY 3079 TCTCTGCAACCGGCACTGTGAAGCTTAAACCTGTGTGCAAGTAATCTTCAAGTCCAAAGA 3138
DB 3241 TCTCTGCAACCGGCACTGTGAAGCTTAAACCTGTGTGCAAGTAATCTTCAAGTCCAAAGA 3300
QY 3139 ATGATGAAGCTGTGTGCGCTTGTGCGCTTCCCAAGCTTAACCTTAACAATTAATGGGCTG 3198
DB 3301 ATGATGAAGCTGTGTGCGCTTGTGCGCTTCCCAAGCTTAACCTTAACAATTAATGGGCTG 3360
QY 3199 TGAAGATGGCAGTACCTGTGTGAATAAAGAAAGCTGTGTGAGAGAGTCAAGTCACTCAAG 3258
DB 3361 TGAAGATGGCAGTACCTGTGTGAATAAAGAAAGCTGTGTGAGAGAGTCAAGTCACTCAAG 3420
QY 3259 CCCCAGTCTGTAATGAGCGTATGTTGGCATCTTTTGAATGAAGATGACCGACAC 3312
DB 3421 CCCCAGTCTGTAATGAGCGTATGTTGGCATCTTTTGAATGAAGATGACCGGATAC 3474

```

RESULT 7  
ADA45218  
ID ADA45218 standard; DNA; 3926 BP.  
XX

```

AC ADA45218;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human MATER DNA.
XX
KW human ; MATER; maternal antigen that embryos require; NTP-ase;
KW nucleoside triphosphate-ase; apoptosis; antifertility; contraceptive;
KW inflammatory; immunosuppressive; gynaecological; endometriosis;
KW ovarian dysfunction; autoimmune premature ovarian failure;
KW autoimmune disease; contraceptive; autoantibody; ds; gene.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT CDS 1..3489
FT /tag=a
FT /product="MATER"
XX
PN EPI285964-A2.
XX
PD 26-FEB-2003.
XX
PE 12-JUL-2002; 2002EP-00090246.
XX
PR 10-AUG-2001; 2001DE-01039874.
XX
PA (SCHD ) SCHERING AG.
XX
PI Weiss B, Lessl M, Peters-Kottig M, Beckmann G;
XX WPI; 2003-302814/30.
XX DR P-PSDB; ADA45219.
XX
PT New nucleic acid for human maternal antigen that embryos require protein,
PT useful e.g. in diagnosis and treatment of female infertility.
XX
PS Claim 1; Page 9-11; 31pp; German.
XX
CC This invention describes a novel human MATER (maternal antigen that
CC embryos require) protein. MATER is an NTP(nucleoside triphosphate)-ase
CC associated with apoptosis, a defect in MATER activity causes growth
CC arrest at the 2-cell stage of fertilized eggs. mRNA transcribed from the
CC MATER gene is found predominantly in ovary, testis and placenta. The
CC products of the invention have antifertility, contraceptive,
CC antiinflammatory, immunosuppressive and gynaecological activity.
CC Effectors of the MATER polypeptide are used to treat infertility.
CC associated with endometriosis and also ovarian dysfunction, autoimmune
CC premature ovarian failure, inflammation, autoimmune diseases and female
CC infertility, and as contraceptives. Measurements of autoantibodies
CC against MATER in body samples or MATER protein or mRNA in eggs is useful
CC for diagnosis of female infertility, particularly by detecting mutations
CC in the MATER gene with a DNA chip. This sequence encodes the human MATER
CC polypeptide described in the invention.
XX
SQ Sequence 3926 BP; 969 A; 958 C; 1061 G; 938 T; 0 U; 0 Other;

```

Query Match 77.7%; Score 3133.6; DB 8; Length 3926;  
Best Local Similarity 95.2%; Pred. NO. 0;  
Matches 3308; Conservative 0; Mismatches 4; Indels 162; Gaps 1;

```

QY 1 ATGGAAGAGAGCAATGCTCACCTTTCCAGTACGGGCTGCAATGCTGCTATGAG 60
DB 1 ATGGAAGAGAGCAATGCTCACCTTTCCAGTACGGGCTGCAATGCTGCTATGAG 60
QY 61 CTAGACAAGAGAAATTCAGACATTCAGAGAAATTAAGAGAAATCTTCAGAAATCG 120
DB 61 CTAGACAAGAGAAATTCAGACATTCAGAGAAATTAAGAGAAATCTTCAGAAATCG 120
QY 121 ACCAGATGCTTATTCACAGATTGAAGTGAAGATGCCAAGTGTGAGTCTGCACTC 180
DB 121 ACCAGATGCTTATTCACAGATTGAAGTGAAGATGCCAAGTGTGAGTCTGCACTC 180

```

181 CTCTGATGAGTATATGAGCATCGCTGGCTGAGCTACGCTATGATCTTTGAA 240  
181 CTCTGATGAGTATATGAGCATCGCTGGCTGAGCTACGCTATGATCTTTGAA 240  
241 AACATGAACCTGCGAACCTCTCGAGAGGACGCGATGATGATGAAA----- 289  
241 AACATGAACCTGCGAACCTCTCGAGAGGACGCGATGATGATGAAAATTCCACAGA 300  
290 ----- 289  
301 GATCTGAAGCAACGATGATGACCAAGGACCAAGCAAGAAAAGTGCCAGAAAATAA 360  
290 ----- 289  
361 TATGCGATGACTAAGCTTATCTTGGGGGTCTGACATCTGACTCGAATATATAAC 420  
290 ----- 318  
421 AAGTATGTTGAATTCATTCTTTTGAGAAATTTCAAGCTATGGAACAAGAT 480  
319 GCCACAGCAGCAGACAGAGAGACAGAAATTTCAAGCTATGGAACAAGATGCTC 378  
481 GCCACAGCAGCAGACAGAGAGACAGAAATTTCAAGCTATGGAACAAGATGCTC 540  
379 ACAGCAGCAGCAGACAGAGAGACAGAAATTTGAGAGTGAACATGGAAGCTACAAGATCAC 438  
541 ACAGCAGCAGCAGACAGAGAGACAGAAATTTGAGAGTGAACATGGAAGCTACAAGATCAC 600  
439 GTGATGACCAATTTGCTGAGAGAGAGATGATCCTGATGTTTGAACAACTGCTCT 498  
601 GTGATGACCAATTTGCTGAGAGAGAGATGATCCTGATGTTTGAACAACTGCTCT 660  
499 GACTGGCCGGAATGCAAACTGTTGCTGCTGTTTGAATTCAGACCGGTGGGCTTCCG 558  
661 GACTGGCCGGAATGCAAACTGTTGCTGCTGTTTGAATTCAGACCGGTGGGCTTCCG 720  
559 CCTGCGAGGTGCTCTGACGGAAGTCAAGAAATTTGGGAAATTCGGCTTACCGCAAGAG 618  
721 CCTGCGAGGTGCTCTGACGGAAGTCAAGAAATTTGGGAAATTCGGCTTACCGCAAGAG 780  
619 ATCTGCTGTGCTGGGCGCAAGTGAATCTTACAGGAGATTTCTCTCACTCTTCTTC 678  
781 ATCTGCTGTGCTGGGCGCAAGTGAATCTTACAGGAGATTTCTCTCACTCTTCTTC 840  
679 CTCTCCGTTAGAGATGACGCGGAAGAGAGACAGTGTCAAGAGTTCATCTCCAG 738  
841 CTCTCCGTTAGAGATGACGCGGAAGAGAGACAGTGTCAAGAGTTCATCTCCAG 900  
739 GAGTGGCCAGATCTCCAGGCTCCGCTGACGAGATCATGTCCGACCAAGAAAGGCTTTG 798  
901 GAGTGGCCAGATCTCCAGGCTCCGCTGACGAGATCATGTCCGACCAAGAAAGGCTTTG 960  
799 TTTCATCATTTGACGTTTGCATGACCTGAGGCTGTCTCTCAACATGACCAAGAGCTTGC 858  
961 TTTCATCATTTGACGTTTGCATGACCTGAGGCTGTCTCTCAACATGACCAAGAGCTTGC 1020  
859 AAAAGCTGAGGCTGAGAGACGCTTCGTTCAACCTCATACGAGTCTGCTGAGAGGTC 918  
1021 AAAAGCTGAGGCTGAGAGACGCTTCGTTCAACCTCATACGAGTCTGCTGAGAGGTC 1080  
919 CTGCTCCCTGAGTCTTCTCTGATGCTCACTGACAGAGAGCTGAGGCAAGAGAGCTCAAG 978  
1081 CTGCTCCCTGAGTCTTCTCTGATGCTCACTGACAGAGAGCTGAGGCAAGAGAGCTCAAG 1140  
979 TCAGAGTGTGCTCTCCGCTTACCTGTTAGTTAGAGAGATCTCCGAGGAAACAAAGATC 1038  
1141 TCAGAGTGTGCTCTCCGCTTACCTGTTAGTTAGAGAGATCTCCGAGGAAACAAAGATC 1200  
1039 CACTTGTCTTTGAGCGCGGAGTTGATGACATCAAGAGACCAAGGGTTGCTGCTGATC 1098  
1201 CACTTGTCTTTGAGCGCGGAGTTGATGATGATGACATCAAGAGACCAAGGGTTGCTGCTGATC 1260  
1099 ATGAACAACCGTGAAGTGTCTGACCAAGTGCAGAGTGCCTCCGCTGCTCTCATCTGC 1158

1261 ATGAACAACCGTGAAGTGTCTGACCAAGTGCAGAGTGCCTCCGCTGCTCTCATCTGC 1320  
1159 GTGGCCCTGACGCTGACAGACGCTGTGGGAGAGAGCGTGCCTCTTCAACCAACGCTC 1218  
1321 GTGGCCCTGACGCTGACAGACGCTGTGGGAGAGAGCGTGCCTCTTCAACCAACGCTC 1380  
1219 ACAAGCTGACAGCGCTTTTGTGTTCATGAGTCAACCTCTGAGGCTGTGCTCCGGTGC 1278  
1381 ACAAGCTGACAGCGCTTTTGTGTTCATGAGTCAACCTCTGAGGCTGTGCTCCGGTGC 1440  
1279 TGTCTCAATCTGAGAGAAAGTGTCTGAGAGGCTTCTGCTGCTGATGAGTGTGAGAGGA 1338  
1441 TGTCTCAATCTGAGAGAAAGTGTCTGAGAGGCTTCTGCTGCTGATGAGTGTGAGAGGA 1500  
1339 GTTGAATTAAGAACTGATGTTTGAATGAGACGCTCATAGTTCAAGATCTCGGAG 1398  
1501 GTTGAATTAAGAACTGATGTTTGAATGAGACGCTCATAGTTCAAGATCTCGGAG 1560  
1399 TCTGAGCTCGTGTCTGTTCATGATGAACATCTTCTCCAGACGCTGATGAGAG 1458  
1561 TCTGAGCTCGTGTCTGTTCATGATGAACATCTTCTCCAGACGCTGATGAGAG 1620  
1459 TACTACACCTTCTTCACTGATCTCCAGGACCTTGTGCGGCTTGTACTACGTTT 1518  
1621 TACTACACCTTCTTCACTGATCTCCAGGACCTTGTGCGGCTTGTACTACGTTT 1680  
1519 GAGGCTGAGAAATGACAGCAGCTCTGCTGCTGATGATGAGAGACCAAGAGTGC 1578  
1681 GAGGCTGAGAAATGACAGCAGCTCTGCTGCTGATGATGAGAGACCAAGAGTGC 1740  
1579 ATGAGCTTAAACAGGACGCTTCTCATATCACTGCTTGTGATGAGAGCTTCTTCTT 1638  
1741 ATGAGCTTAAACAGGACGCTTCTCATATCACTGCTTGTGATGAGAGCTTCTTCTT 1800  
1639 GAGCTGAG 1698  
1801 GAGCTGAG 1860  
1699 CTGGGGGTGAAGCAGAACTTCTGACAGTGGTCTCTGTTGGGTCAAGACCTTAATGC 1758  
1861 CTGGGGGTGAAGCAGAACTTCTGACAGTGGTCTCTGTTGGGTCAAGACCTTAATGC 1920  
1759 ACCACCCAGAGAGACCTTGAAGGCTTCACTGTCTTTTGAAGACTCAAGACCAAGAG 1818  
1921 ACCACCCAGAGAGACCTTGAAGGCTTCACTGTCTTTTGAAGACTCAAGACCAAGAG 1980  
1819 TTTGTTGCTTGGCATTTAAACAGCTTCCAAAGAGTGGGCTTCCGATTAACAGAGAGCTG 1878  
1981 TTTGTTGCTTGGCATTTAAACAGCTTCCAAAGAGTGGGCTTCCGATTAACAGAGAGCTG 2040  
1879 GACTTGAATGATCTCTCTTCTGCTCCAGACAGTCCGATTTTGGGAGAAATTCGAGTG 1938  
2041 GACTTGAATGATCTCTCTTCTGCTCCAGACAGTCCGATTTTGGGAGAAATTCGAGTG 2100  
1939 GATGTCAAGAGATCTTCCAAAGATGAGTCCGCTGAGGAGTCTGTGTCTCTTA 1998  
2101 GATGTCAAGAGATCTTCCAAAGATGAGTCCGCTGAGGAGTCTGTGTCTCTTA 2160  
1999 TGGATGCGGGAATTAAGACCTCATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2058  
2161 TGGATGCGGGAATTAAGACCTCATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220  
2059 ACCACCCAGACCTGCGGAG 2118  
2221 ACCACCCAGACCTGCGGAG 2280  
2119 AAGACCTGTGTGCAAGCTGAGGATCTCCAGCTTGAAGATTAAGAGAGAGAGAGAGAG 2178  
2281 AAGACCTGTGTGCAAGCTGAGGATCTCCAGCTTGAAGATTAAGAGAGAGAGAGAGAG 2340  
2179 AATGACAGAGATTAACCTCTGAGTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2238

```

Db 2341 AATGCAGATTATACCTCTGTGTGACGACCTCTGAGAAATGTCATGCGCAACCGTAAAC 2400
Qy 2239 CTAAGATCCCTCAACTTGGAGGACCCACCTGAAAGAAAGGATTAAGGATGGCGTGT 2298
Db 2401 CTAAGATCCCTCAACTTGGAGGACCCACCTGAAAGAAAGGATTAAGGATGGCGTGT 2460
Qy 2299 GAAGCCTTAAACACCCAAAATGTTGTGTGAGTCTTGTGAGGCTGGATTGTGCTGTGATTTG 2358
Db 2461 GAAGCCTTAAACACCCAAAATGTTGTGTGAGTCTTGTGAGGCTGGATTGTGCTGTGATTTG 2520
Qy 2359 ACCCATGCTGTACTTGAAGATCTCCCAATCCTTACGACCTCCCAAGCTCTGAATCT 2418
Db 2521 ACCCATGCTGTACTTGAAGATCTCCCAATCCTTACGACCTCCCAAGCTCTGAATCT 2580
Qy 2419 CTGAGCCTGGAGAGAAACAGAGTGAAGACCAAGGAGTAAATGCTCTGAGTAAAGCTTTG 2478
Db 2581 CTGAGCCTGGAGAGAAACAGAGTGAAGACCAAGGAGTAAATGCTCTGAGTAAAGCTTTG 2640
Qy 2479 AGAGTCTCCAGATGCGCCCTGCGAGAAAGCTGATATCTGAGAGACTGTGTGATGACAGCCACG 2538
Db 2641 AGAGTCTCCAGATGCGCCCTGCGAGAAAGCTGATATCTGAGAGACTGTGTGATGACAGCCACG 2700
Qy 2539 GGTGTCCAGAGTCTGTGGCTCAGCCCTGTGACGAAACCGGAGCTTGAACACCTGTGCTTA 2598
Db 2701 GGTGTCCAGAGTCTGTGGCTCAGCCCTGTGACGAAACCGGAGCTTGAACACCTGTGCTTA 2760
Qy 2599 TCACAACACAGCTGTGGAGAGAGAGTGAATCTACTGTGTGATTCATGAGGCTTCC 2658
Db 2761 TCACAACACAGCTGTGGAGAGAGAGTGAATCTACTGTGTGATTCATGAGGCTTCC 2820
Qy 2659 CACTGTACTGTGACAGAGGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2718
Db 2821 CACTGTACTGTGACAGAGGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2880
Qy 2719 TTTCTTGACATTTGGCTTATGGGTATCTATGAGTCTGAGGACCTGAGGCTTATGATGATGATGAT 2778
Db 2881 TTTCTTGACATTTGGCTTATGGGTATCTATGAGTCTGAGGACCTGAGGCTTATGATGATGATGAT 2940
Qy 2779 CCTGTGAAGAACAATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2838
Db 2941 CCTGTGAAGAACAATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
Qy 2839 CTCAGAGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2898
Db 3001 CTCAGAGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060
Qy 2899 TGTGTGATCTGAGAGAGACACCTGAAAGGCTGTGATCTCAAGGACCAATGCTTGGGT 2958
Db 3061 TGTGTGATCTGAGAGAGACACCTGAAAGGCTGTGATCTCAAGGACCAATGCTTGGGT 3120
Qy 2959 GACGGTGGGTTGCTGCGCTGTGCGAGGAGACTGAAAGCAAAAGACAGTCTTGAAGAGAGAGAG 3018
Db 3121 GACGGTGGGTTGCTGCGCTGTGCGAGGAGACTGAAAGCAAAAGACAGTCTTGAAGAGAGAGAG 3180
Qy 3019 CTCGGGTTGAAGGACATGTAAGTGAAGTCTTGAATGCTGTGAGGACCTCTCTTGGCCCTT 3078
Db 3181 CTCGGGTTGAAGGACATGTAAGTGAAGTCTTGAATGCTGTGAGGACCTCTCTCTTGGCCCTT 3240
Qy 3079 TCTGTGACCAACCGGACATCTTCAACAGTCTTAAACCTGTGTGAGGAGTAACTTGAATGCCAAAG 3138
Db 3241 TCTGTGACCAACCGGACATCTTCAACAGTCTTAAACCTGTGTGAGGAGTAACTTGAATGCCAAAG 3300
Qy 3139 ATGATGAAGCTGTGTGTGGGCTTTTGGCTGTGCTGCAAGTCTTAACTTGAATTTGGGCTG 3198
Db 3301 ATGATGAAGCTGTGTGTGGGCTTTTGGCTGTGCTGCAAGTCTTAACTTGAATTTGGGCTG 3360
Qy 3199 TGAAGATGAGAGTACCTGTGCAAAATGAAGAGCTGTGAGAGAGTGCAGTACTCAAG 3258
Db 3361 TGAAGATGAGAGTACCTGTGCAAAATGAAGAGCTGTGAGAGAGTGCAGTACTCAAG 3420
Qy 3259 CCCCAGTCTGTATTTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3312
Db 3421 CCCCAGTCTGTATTTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3474

```

```

RESULT 8
ID AAL47135 standard; DNA; 5859 BP.
XX AAL47135;
AC AAL47135;
DT 20-AUG-2002 (first entry)
DE Pyrin domain containing protein NALP8/Py12 coding sequence.
XX Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
XX antiarteriosclerotic; antipsoriatic; antibacterial; virocidic;
XX neuroprotective; antiarthritic; antirheumatic; antiaesthetic;
XX nephrotropic; osteopathic; nootropic; intracellular signal transduction;
XX inflammation; Alzheimer's disease; infection; psoriasis; asthma;
XX arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
XX osteoarthritis; glomerulonephritis; gene; ds.
XX Unidentified.
XX MO200240668-A2.
XX 23-MAY-2002.
XX 30-OCT-2001; 2001WO-EP012545.
XX 15-NOV-2000; 2000DE-01056687.
XX 30-NOV-2000; 2000DE-01059595.
XX (APOT-) APOTECHE RES & DEV LTD.
XX Tschoep J, Martison F;
XX WPI; 2002-427093/45.
XX P-PSDB; MA017863.
XX New DNA encoding protein with pyrin domain, useful for treating diseases
XX PT involving impaired signal transduction, particularly inflammation, also
XX PT proteins and antibodies.
XX PS Claim 5; Fig 1; 116pp; German.
XX CC The present invention relates the DNA and their encoded proteins, where
XX CC the proteins contain at least one PYD (pyrin) domain. These can be used
XX CC to treat diseases associated with impaired intracellular signal
XX CC transduction, particularly inflammation such as psoriasis,
XX CC arteriosclerosis, bacterial or viral infections (particularly meningitis
XX CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
XX CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
XX CC and Parkinson's diseases. The present sequence is a coding sequence of
XX CC the invention
XX SQ Sequence 5859 BP; 1512 A; 1443 C; 1530 G; 1374 T; 0 U; 0 Other;
Query Match 75.8%; Score 3057.8; DB 6; Length 5859;
Best Local Similarity 96.2%; Pred. No. 0; Mismatches 2; Indels 123; Gaps 1;
Matches 3192; Conservative 0;
Qy 1 ATGAGAGAGAGCAAAATGCTACCTTTTCCAGCTACGAGGCTGCAATGCTGTCTATGAG 60
Db 1975 ATGAGAGAGAGCAAAATGCTACCTTTTCCAGCTACGAGGCTGCAATGCTGTCTATGAG 2034
Qy 61 CTGACCAAGAGAGATTTGACACATTTCAAGGATTTACTAAAGAGAAATCTTCAGAAATCG 120
Db 2035 CTGACCAAGAGAGATTTGACACATTTCAAGGATTTACTAAAGAGAAATCTTCAGAAATCG 2094
Qy 121 ACCAGATGCTCTATTTCCAGAGTTTGAATGAGAGATGCCAAGGAGATGCTGGCACTC 180
Db 2095 ACCAGATGCTCTATTTCCAGAGTTTGAATGAGAGATGCCAAGGAGATGCTGGCACTC 2154
Qy 181 CTCTTGATGAGATTTATGAGAGATCGCTGGCTGTGGCTACGTCATTAAGATCTTTGAA 240

```



```

Db 2155 CTTTGATGATGATATTATGAGCATGCGGCTGGGCTAGCTCATTTAGATCTTTGAA 2214
Qy 241 AACATGAACCTGCGAACTCTTCGAGAAAGGCACGGATGACATGAAAA----- 289
Db 2215 AACATGAACCTGCGAACTCTTCGAGAAAGGCACGGATGACATGAAAAAGATTCACCA 2274
Qy 290 ----- 289
Db 2275 GAAGATCCTGAAGCAAGATGACTGACCAAGGACCAAGCAAGAAAAAGTCCAGGAATT 2334
Qy 290 -----AAATTTCA 297
Db 2335 TCACAAGCTGTGCAACAAGATAGTGCAACGCTGCAAGAGCAAAAGAACAAAGAAATTTCA 2394
Qy 298 CAAGCTATGAAACAAGAGGTGCCACAGACAGACAGAGAAACAAGAAATTTTCA 357
Db 2395 CAAGCTATGAAACAAGAGGTGCCACAGACAGACAGAGAAACAAGAAATTTTCA 2454
Qy 358 GCTATGAAACAAGAGGTGCCACAGACAGAGAAACAAGAAACAAGAAATTTGAGTAC 417
Db 2455 GCTATGAAACAAGAGGTGCCACAGACAGAGAAACAAGAAACAAGAAATTTGAGTAC 2514
Qy 418 ACATGGAATCAAGAGTCAAGTATGACCAAAATTCGCTGAGAGAGAGATGATGCTGCT 477
Db 2515 ACATGGAATCAAGAGTCAAGTATGACCAAAATTCGCTGAGAGAGAGATGATGCTGCT 2574
Qy 478 AGTTTGAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537
Db 2575 AGTTTGAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2634
Qy 538 TCAGACCGGTGGGCTTCCGCTGCGACAGGTGCTTTCGACAGGAAATTCAGAAATTTGGG 597
Db 2635 TCAGACCGGTGGGCTTCCGCTGCGACAGGTGCTTTCGACAGGAAATTCAGAAATTTGGG 2694
Qy 598 AAATGCGCTCTGACCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
Db 2695 AAATGCGCTCTGACCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2754
Qy 658 ATGTTCTCTTAAGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 717
Db 2755 ATGTTCTCTTAAGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2814
Qy 718 GTCAAGAGTTCATCTCCAGGAGTGGGCAAGCTCCAGGCTCCGCTGACAGAGATCATG 777
Db 2815 GTCAAGAGTTCATCTCCAGGAGTGGGCAAGCTCCAGGCTCCGCTGACAGAGATCATG 2874
Qy 778 TCCGACCAAGAAAGCTGTTGTTTCAATGACGGTTTCGATGACCTGGGCTCTGCTCTC 837
Db 2875 TCCGACCAAGAAAGCTGTTGTTTCAATGACGGTTTCGATGACCTGGGCTCTGCTCTC 2934
Qy 838 AACATGACACAAAGCTCTGCAAGAGTGGGCTGAGAAAGCAAGCTCCGTTCACTCTCAT 897
Db 2935 AACATGACACAAAGCTCTGCAAGAGTGGGCTGAGAAAGCAAGCTCCGTTCACTCTCAT 2994
Qy 898 CGCAATCTGCTGAGAGAGTCTGCTCTCTGAGTCTTCTCTGATGTCACGCTCAGAGAC 957
Db 2995 CGCAATCTGCTGAGAGAGTCTGCTCTCTGAGTCTTCTCTGATGTCACGCTCAGAGAC 3054
Qy 958 GTGGGCAAGAGAGTCAAGTCAAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1017
Db 3055 GTGGGCAAGAGAGTCAAGTCAAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3114
Qy 1018 ATCTCCGGGGAACAAGATTCATCTGCTCTTGAAGCGGAGATGTTGAGCATCAGAG 1077
Db 3115 ATCTCCGGGGAACAAGATTCATCTGCTCTTGAAGCGGAGATGTTGAGCATCAGAG 3174
Qy 1078 ACAAGAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137
Db 3175 ACAAGAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3234
Qy 1138 GCGGAGGCTCTCTCATGCGGTGGCCCTGACGCTGACAGAGATGTTGGGAGAGCGTC 1197

```

```

Db 3235 GCGGAGGCTCTCTCATCTGCGGTGGCCCTGACGCTGACAGAGTGTGGGAGAGCGTC 3294
Qy 1198 GCGGCTTTCAACCAAGCTCTCAAGGCTTGACAGCGCGCTTTTGTGTTTCACTGACCC 1257
Db 3295 GCGGCTTTCAACCAAGCTCTCAAGGCTTGACAGCGCGCTTTTGTGTTTCACTGACCC 3354
Qy 1258 CTTGAGGCTGTGTCGCGGCTGCTCTCAATCTGAGAGAAAGATGTTCTGAAAGCGCTTC 1317
Db 3355 CTTGAGGCTGTGTCGCGGCTGCTCTCAATCTGAGAGAAAGATGTTCTGAAAGCGCTTC 3414
Qy 1318 TGCCGTATGCTGTGAGAGAGTGTGAATAAGAGTCAAGTGTGATGAGAGAGCTC 1377
Db 3415 TGCCGTATGCTGTGAGAGAGTGTGAATAAGAGTCAAGTGTGATGAGAGAGCTC 3474
Qy 1378 ATGTTCAAGAGCTGCGGAGATCTGAGCTCCGCTCTGTTTCAATGAACATCTTCTC 1437
Db 3475 ATGTTCAAGAGCTGCGGAGATCTGAGCTCCGCTCTGTTTCAATGAACATCTTCTC 3534
Qy 1438 CAGACAGCTGAGAGATGACACCTTTTCAACCTCAGTCTCAGAGCTTCTGCT 1497
Db 3535 CAGACAGCTGAGAGATGACACCTTTTCAACCTCAGTCTCAGAGCTTCTGCT 3594
Qy 1498 GCGGCTTTGATCACTGCTGCTGAGAGGCTGGAATGAGCAGGCTCTGCTCTGATC 1557
Db 3595 GCGGCTTTGATCACTGCTGCTGAGAGGCTGGAATGAGCAGGCTCTGCTCTGATC 3654
Qy 1558 GTTGAAGAGCAAGAGTCCATGAGCTTAAACAGCAGGCTTCCATATCACTGCTT 1617
Db 3655 GTTGAAGAGCAAGAGTCCATGAGCTTAAACAGCAGGCTTCCATATCACTGCTT 3714
Qy 1618 TGAATGAAGCGTTTCTGTTTGGCTGCTGAGAGCAAGAGTGAAGAGGCACTGAGGTC 1677
Db 3715 TGAATGAAGCGTTTCTGTTTGGCTGCTGAGAGCAAGAGTGAAGAGGCACTGAGGTC 3774
Qy 1678 CTGCTGAGCTGCTCCGTTTCCCTGAGGAGTGAAGAGCTTCTGACCTGAGGCTCTCTG 1737
Db 3775 CTGCTGAGCTGCTCCGTTTCCCTGAGGAGTGAAGAGCTTCTGACCTGAGGCTCTCTG 3834
Qy 1738 TTGGGTGAGCAGCTTAATGCAACACCCAGAGACACCTGAGAGGCTTCACTGCTT 1797
Db 3835 TTGGGTGAGCAGCTTAATGCAACACCCAGAGACACCTGAGAGGCTTCACTGCTT 3854
Qy 1798 TTGAGAGTCAAGCAAGAGTGTGCTGCTGCAATTAACACCTTCCAAAGAGTGG 1857
Db 3895 TTGAGAGTCAAGCAAGAGTGTGCTGCTGCAATTAACACCTTCCAAAGAGTGG 3954
Qy 1858 CTTCCGATTAACCAAGCTGGAATGATAGATCTTCTCTCTCTCTCTCTCTCTCTCTCT 1917
Db 3955 CTTCCGATTAACCAAGCTGGAATGATAGATCTTCTCTCTCTCTCTCTCTCTCTCTCT 4014
Qy 1918 TATTTGGGAAATTTGGGTGATGTCAAGAGATCTTCCAGAGATGATGCTGCTGAG 1977
Db 4015 TATTTGGGAAATTTGGGTGATGTCAAGAGATCTTCCAGAGATGATGCTGCTGAG 4074
Qy 1978 GCATGCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2037
Db 4075 GCATGCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4134
Qy 2038 GATTTCTGCTCAGCTTGGACCAACCAACCTGAGGAGCTGAGAGCTGAGGAGCAGC 2097
Db 4135 GATTTCTGCTCAGCTTGGACCAACCAACCTGAGGAGCTGAGAGCTGAGGAGCAGC 4194
Qy 2098 ATCTGACAGAGCGGACATGAGAGCCTGTGTGCAAGCTGAGGATCCCACTGCAAG 2157
Db 4195 ATCTGACAGAGCGGACATGAGAGCCTGTGTGCAAGCTGAGGATCCCACTGCAAG 4254
Qy 2158 ATACAGAGCTGATGTTTGAAGATGACAGATTAACCTCTGTGTGACAGCACTTGGAGA 2217
Db 4255 ATACAGAGCTGATGTTTGAAGATGACAGATTAACCTCTGTGTGACAGCACTTGGAGA 4314
Qy 2218 ATGCTCATGAGCAACCGGATCAAGATCCCTCAACTTGGAGGACCAACCTGAAGAA 2277
Db 4315 ATGCTCATGAGCAACCGGATCAAGATCCCTCAACTTGGAGGACCAACCTGAAGAA 4374

```

```

OY 2278 GAGATGTAAGATGGCGGTGTAAGCCTTAAACACCCAAAATGTTTGTAAGTCTTTG 2337
    |||||||
DB 4375 GAGATGTAAGATGGCGGTGTAAGCCTTAAACACCCAAAATGTTTGTAAGTCTTTG 4434
OY 2338 AGCTGTGATGCTGTGATTTGACCCATGCTGTGTAACCTGAAGATCTCCAAATCTTACG 2397
    |||||||
DB 4435 AGCTGTGATGCTGTGATTTGACCCATGCTGTGTAACCTGAAGATCTCCAAATCTTACG 4494
OY 2398 ACCCTCCCAAGCCTTAATCTCTGAGCCTGAGAGAAAAGATGACAGACCAAGGAATA 2457
    |||||||
DB 4495 ACCCTCCCAAGCCTTAATCTCTGAGCCTGAGAGAAAAGATGACAGACCAAGGAATA 4554
OY 2458 ATGCTCTCAGTATGCTGTGAGAGTCTCCAGTGCAGCCTGACAGAGCTGATCTGAG 2517
    |||||||
DB 4555 ATGCTCTCAGTATGCTGTGAGAGTCTCCAGTGCAGCCTGACAGAGCTGATCTGAG 4614
OY 2518 GACTGTGATCATCACAGCCAGGGTTGCGAGAGTCTGAGCTCAGCCCTGTCAGCAACCGG 2577
    |||||||
DB 4615 GACTGTGATCATCACAGCCAGGGTTGCGAGAGTCTGAGCTCAGCCCTGTCAGCAACCGG 4674
OY 2578 AGCTTGACACACTGTGCTTATCCAAACAAGCTGTGGGAAACGAAGGTGTAAATCTTACTG 2637
    |||||||
DB 4675 AGCTTGACACACTGTGCTTATCCAAACAAGCTGTGGGAAACGAAGGTGTAAATCTTACTG 4734
OY 2638 TGTGATCATGAGAGCTTCCCACTGTAGTCTGACAGAGGCTGATGCTGAATCAGTGCAC 2697
    |||||||
DB 4735 TGTGATCATGAGAGCTTCCCACTGTAGTCTGACAGAGGCTGATGCTGAATCAGTGCAC 4794
OY 2698 CTGACACAGCGCTGAGTGTGTTTCTTGCACTTGCGCTTAATGAGTAACTCATGCTGACG 2757
    |||||||
DB 4795 CTGACACAGCGCTGAGTGTGTTTCTTGCACTTGCGCTTAATGAGTAACTCATGCTGACG 4854
OY 2758 CACCTGACCGCTTATGATTAACCTCTGTGAAGACATATGCGCTGAAGCTTCTGTGCAAGTTC 2817
    |||||||
DB 4855 CACCTGACCGCTTATGATTAACCTCTGTGAAGACATATGCGCTGAAGCTTCTGTGCAAGTTC 4914
OY 2818 ATGAGAGAACAATCTTGTATCTCCAGAGACCTGAGTGTGTAAGTGTATCTATCTCACCCGC 2877
    |||||||
DB 4915 ATGAGAGAACAATCTTGTATCTCCAGAGACCTGAGTGTGTAAGTGTATCTATCTCACCCGC 4974
OY 2878 GCGTCTGTGAGAGTCTGTCTCTGTGTGATCTTGAGAGACAGACACTGAAGAAGCTGTGAT 2937
    |||||||
DB 4975 GCGTCTGTGAGAGTCTGTCTCTGTGTGATCTTGAGAGACAGACACTGAAGAAGCTGTGAT 5034
OY 2938 CTACAGGACAATGCGCTGTGATGACGTGGGTGCTGCGCTGTGTGAGAGGACTGAAGCAA 2997
    |||||||
DB 5035 CTACAGGACAATGCGCTGTGATGACGTGGGTGCTGCGCTGTGTGAGAGGACTGAAGCAA 5094
OY 2998 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGGACATGAGACTGATCTTGATTTGCTGT 3057
    |||||||
DB 5095 AAGAACAGTGTCTGACGAGACTCGGGTTGAAGGACATGAGACTGATCTTGATTTGCTGT 5154
OY 3058 GAGGCACTCTCTTGGCCCTTCTCTGCAACCGGCACTGACCAAGTCTTAAACCTGTGTGAG 3117
    |||||||
DB 5155 GAGGCACTCTCTTGGCCCTTCTCTGCAACCGGCACTGACCAAGTCTTAAACCTGTGTGAG 5214
OY 3118 AATACTTCACTGCTCCAAAGAAATGATGAAGCTGTGTGAGCCTTTGCTGTGCCACGCTCT 3177
    |||||||
DB 5215 AATACTTCACTGCTCCAAAGAAATGATGAAGCTGTGTGAGCCTTTGCTGTGCCACGCTCT 5274
OY 3178 AACTTACAGATTAATTTGG 3194
    |||||||
DB 5275 AACTTACAGATTAATTTGG 5291
    |||||||

```

## RESULT 9

AA147131 standard; DNA; 6939 BP.

AA147131;

20-AUG-2002 (first entry)

```

XX DE Pylrin domain containing protein NALPs/Pyls-hs coding sequence.
XX KM Pylrin domain; PYD domain; antiinflammatory; antiparkinsonian;
XX KM antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
XX KM neuroprotective; antiallergic; antirheumatic; antiaesthetic;
XX KM nephrotoxic; osteopathic; neotropic; intracellular signal transduction;
XX KM inflammation; Alzheimer's disease; infection; psoriasis; asthma;
XX KM arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
XX KM osteoarthritis; glomerulonephritis; gene; ds.
XX OS Unidentified.
XX PN MO200240668-A2.
XX PD 23-MAY-2002.
XX PF 30-OCT-2001; 2001WO-EP012545.
XX PR 15-NOV-2000; 2000DE-01056687.
XX PR 30-NOV-2000; 2000DE-01059595.
XX PA (APOT-) APOTECH RES & DEV LTD.
XX PI Techopp J, Martinson F;
XX XX WPI; 2002-427093/45.
XX DR P-PSDB; AAO17859.
XX PT New DNA encoding protein with pyrin domain, useful for treating diseases
XX PT involving impaired signal transduction, particularly inflammation, also
XX PT proteins and antibodies.
XX PS Claim 5; Fig 1; 116pp; German.
XX CC The present invention relates the DNA and their encoded proteins, where
XX CC the proteins contain at least one PYD (pyrin) domain. These can be used
XX CC to treat diseases associated with impaired intracellular signal
XX CC transduction, particularly inflammation such as psoriasis,
XX CC arteriosclerosis, bacterial or viral infections (particularly meningitis
XX CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
XX CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
XX CC and Parkinson's diseases. The present sequence is a coding sequence of
XX CC the invention
XX SQ Sequence 6939 BP; 1791 A; 1719 C; 1801 G; 1628 T; 0 U; 0 Other;
XX
XX Query Match 75.8%; Score 3057.8; DB 6; Length 6939;
XX Best Local Similarity 96.2%; Pred. No. 0;
XX Matches 3192; Conservative 0; Mismatches 2; Indels 123; Gaps 1;
OY 1 ATGAGAAGAGACAATGCTTCACTTTTCCAGCTACGCGGCTGCAATGCTGTCTTATGAG 60
    |||||||
DB 3055 ATGAGAAGAGACAATGCTTCACTTTTCCAGCTACGCGGCTGCAATGCTGTCTTATGAG 3114
OY 61 CTGACAAGAGAAGATTTCAGACTTCACGAATTAATAAGAAAGAAATCTTCAGAAATCG 120
    |||||||
DB 3115 CTGACAAGAGAAGATTTCAGACTTCAGGAATTAATAAGAAAGAAATCTTCAGAAATCG 3174
OY 121 ACCACATGCTCTATTCACAGATTGGAATGAGAAATGCAACGATGAGATGCTGGCACTC 180
    |||||||
DB 3175 ACCACATGCTCTATTCACAGATTGGAATGAGAAATGCAACGATGAGATGCTGGCACTC 3234
OY 181 CTCTTGATGAGATTTATGAGAGCATGCTGCGCTGTGGCTACGTCATTTAGATCTTTGAA 240
    |||||||
DB 3235 CTCTTGATGAGATTTATGAGAGCATGCTGCGCTGTGGCTACGTCATTTAGATCTTTGAA 3294
OY 241 AACATGAACCTGGACACCTCTCGAGAGAAGACGCGGATGATCATGATAAAGCATTCACA 289
    |||||||
DB 3295 AACATGAACCTGGACACCTCTCGAGAGAAGACGCGGATGATCATGATAAAGCATTCACA 3354
    |||||||
OY 290 ----- 289

```

Db 3355 GAAGATCTGAAGACGATGACTGACCAAGGACCAAGGAAAAATGCGCAGGAATT 3414  
Qy 290 -----AAATTTCA 297  
Db 3415 TCACAACTGTGCAACAAAGTAGTGCCACAGCTGCAGAGACAAAGAACAAAGAAATTTCA 3474  
Qy 298 CAAGCTATGGAACAAAGAGGTGCAACAGAGACAGAGAAAGAAATTTTCAAA 357  
Db 3475 CAAGCTATGGAACAAAGAGGTGCAACAGAGACAGAGAAAGAAATTTTCAAA 3534  
Qy 358 GCTATGGAACAAAGAGGTGCAACAGAGACAGAGAAAGAAATTTTCAAA 417  
Db 3535 GCTATGGAACAAAGAGGTGCAACAGAGACAGAGAAAGAAATTTTCAAA 3594  
Qy 418 ACATGGGATCTACAAAGATGACGTGATGACCAAAATTTGCTGAGAGAGAGATGATGCTGCT 477  
Db 3595 ACATGGGATCTACAAAGATGACGTGATGACCAAAATTTGCTGAGAGAGAGATGATGCTGCT 3654  
Qy 478 AGTTTGAAGAACATGCTGCTGACCTGCGGAAATGCAACGTTGGCTGCTTTGAT 537  
Db 3655 AGTTTGAAGAACATGCTGCTGACCTGCGGAAATGCAACGTTGGCTGCTTTGAT 3714  
Qy 538 TCAGACCGGTGGGCTTCCGCGCTCGCAAGTGTGCTGCAAGAAATGCAAGAAATGGG 597  
Db 3715 TCAGACCGGTGGGCTTCCGCGCTCGCAAGTGTGCTGCAAGAAATGCAAGAAATGGG 3774  
Qy 598 AAATCGGCTCTAGCCGAAAGATCGTGTGTGTGTGGGCGCAAGTGTGATCTTACCGAGGA 657  
Db 3775 AAATCGGCTCTAGCCGAAAGATCGTGTGTGTGTGGGCGCAAGTGTGATCTTACCGAGGA 3834  
Qy 658 ATGTTCTCTAGCTCTTCTCCCGTTAGAGAGATGACAGCGGAAAGAGAGAGCACT 717  
Db 3835 ATGTTCTCTAGCTCTTCTCCCGTTAGAGAGATGACAGCGGAAAGAGAGAGCACT 3894  
Qy 718 GTCAAGAGATTCCTCAAGAGAGTGGCAGACTCCAGAGCTCCGCTGACGAGATCATG 777  
Db 3895 GTCAAGAGATTCCTCAAGAGAGTGGCAGACTCCAGAGCTCCGCTGACGAGATCATG 3954  
Qy 778 TCCCGACCAAGAAAGCTGTTGTTTCAATTCATTCAGCGTTTGCATGACTGTGGCTCTGTCTC 837  
Db 3955 TCCCGACCAAGAAAGCTGTTGTTTCAATTCATTCAGCGTTTGCATGACTGTGGCTCTGTCTC 4014  
Qy 838 AACCAATGACAAAGAGCTGCAAGAGCTGGGCTGAGAGACAGCTCCGTTACCCCTCA 897  
Db 4015 AACCAATGACAAAGAGCTGCAAGAGCTGGGCTGAGAGACAGCTCCGTTACCCCTCA 4074  
Qy 898 CGCAGTCTGCTAGAGAGAGTCTGCTGCTGAGTCCCTGATGCTGACCTGACAGAGC 957  
Db 4075 CGCAGTCTGCTAGAGAGAGTCTGCTGCTGAGTCCCTGATGCTGACCTGACAGAGC 4134  
Qy 958 GTGGGACAGAGAGCTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTTAGTTAGAGA 1017  
Db 4135 GTGGGACAGAGAGCTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTTAGTTAGAGA 4194  
Qy 1018 ATCTCCGGGGAAACAAAGATTCACATTCCTGTTGAGCGGGGATTTGATGATCAAGAG 1077  
Db 4195 ATCTCCGGGGAAACAAAGATTCACATTCCTGTTGAGCGGGGATTTGATGATCAAGAG 4254  
Qy 1078 ACACAAAGGTTGCTGTCGATCATGAACAACGTTGAGTCTGACCAAGTCCAGAGTCCC 1137  
Db 4255 ACACAAAGGTTGCTGTCGATCATGAACAACGTTGAGTCTGACCAAGTCCAGAGTCCC 4314  
Qy 1138 GCGGTGGGCTCTCTCATCTGCGTGGCGCTGCAAGTGCAGAGAGTGGTGGGAGAGAGCTC 1197  
Db 4315 GCGGTGGGCTCTCTCATCTGCGTGGCGCTGCAAGTGCAGAGAGTGGTGGGAGAGAGCTC 4374  
Qy 1198 GCGCGCTTCAACCAACGCTCAACAGGCTGCAAGCGCGCTTTGTGTGTTTATACGCTCAC 1257  
Db 4375 GCGCGCTTCAACCAACGCTCAACAGGCTGCAAGCGCGCTTTGTGTGTTTATACGCTCAC 4434  
Qy 1258 CCTGAGAGGCTGTCCGCGCTGTCTCAATCTGAGAGAAAGATTGCTGTAAGCGCTTC 1317  
Db 4435 CCTGAGAGGCTGTCCGCGCTGTCTCAATCTGAGAGAAAGATTGCTGTAAGCGCTTC 4494

Qy 1318 TGCCGTATGCTGTGAGAGAGTGTGGAATAGGAATGCAAGTGTGATGATGAGAGAGCTC 1377  
Db 4495 TGCCGTATGCTGTGAGAGAGTGTGGAATAGGAATGCAAGTGTGATGAGAGAGAGCTC 4554  
Qy 1378 ATGTTCAAGAGACTGGGAGAGTGTGAGTCCGTGCTGTTTCAATGAACATCTTCTC 1437  
Db 4555 ATGTTCAAGAGACTGGGAGAGTGTGAGTCCGTGCTGTTTCAATGAACATCTTCTC 4614  
Qy 1438 CCAGACAGCACTGTGAGAGAGTACTACCTTCTTCACTCTCATGCTCCAGAGCTTCTGT 1497  
Db 4615 CCAGACAGCACTGTGAGAGAGTACTACCTTCTTCACTCTCATGCTCCAGAGCTTCTGT 4674  
Qy 1498 GCGCGCTTGTACTAGTGTGAGAGAGCTGTGAATAATGAGCCAGCTCTGCTCTGTGAC 1557  
Db 4675 GCGCGCTTGTACTAGTGTGAGAGAGCTGTGAATAATGAGCCAGCTCTGCTCTGTGAC 4734  
Qy 1558 GTTGAAGAGACAAAGAGTCTCATGAGCTTAAACAGCAGAGCTTCATATCCATCGCTT 1617  
Db 4735 GTTGAAGAGACAAAGAGTCTCATGAGCTTAAACAGCAGAGCTTCATATCCATCGCTT 4794  
Qy 1618 TGGATGAAGCGTTTCTTGTGTGCTGTGAGCGAAGCGTAAAGAGGCACTGAGAGTTC 1677  
Db 4795 TGGATGAAGCGTTTCTTGTGTGCTGTGAGCGAAGCGTAAAGAGGCACTGAGAGTTC 4854  
Qy 1678 CTGCTGGGCTGTCCCGTCCCTGCGGGTGAAGAGAGAGCTTTCGACTGTGGTCTCTCG 1737  
Db 4855 CTGCTGGGCTGTCCCGTCCCTGCGGGTGAAGAGAGAGCTTTCGACTGTGGTCTCTCG 4914  
Qy 1738 TTGGGTGAGAGAGCTTAATGCCACCACTCCAGAGAGACCTTGCATGCTCTT 1797  
Db 4915 TTGGGTGAGAGAGCTTAATGCCACCACTCCAGAGAGAGCTTGCATGCTCTT 4974  
Qy 1798 TTGAGACTCAAGACAAAGAGTGTGCTGCTTGCATTAAACAGCTTCAAGAGTGTG 1857  
Db 4975 TTGAGACTCAAGACAAAGAGTGTGCTGCTTGCATTAAACAGCTTCAAGAGTGTG 5034  
Qy 1858 CTTCGATTAACCAAGACCTGAGCTTGTATGATCTTCTCTGCTCAGACCTGTCCG 1917  
Db 5035 CTTCGATTAACCAAGACCTGAGCTTGTATGATCTTCTCTGCTCAGACCTGTCCG 5094  
Qy 1918 TATTTGCGGAAATTCGGGTGATGTCAAGAGATCTTCCCAAGATGAGTCCGCTGAG 1977  
Db 5095 TATTTGCGGAAATTCGGGTGATGTCAAGAGATCTTCCCAAGATGAGTCCGCTGAG 5154  
Qy 1978 GCATGCTCTGTGTCTCTCTATGATGCGGATTAAGACCTCATTTGAGAGACAGTGGAA 2037  
Db 5155 GCATGCTCTGTGTCTCTCTATGATGCGGATTAAGACCTCATTTGAGAGACAGTGGAA 5214  
Qy 2038 GATTTCTGCTCAATGCTTTGGACCCACCACTGCGGAGCTGAGCTGGGACAGC 2097  
Db 5215 GATTTCTGCTCAATGCTTTGGACCCACCACTGCGGAGCTGAGCTGGGACAGC 5274  
Qy 2098 ATTCGACAGAGCGGGACCATGAAGACCTGTGTGCAAGCTGAGAGCATCCACCTGCAAG 2157  
Db 5275 ATTCGACAGAGCGGGACCATGAAGACCTGTGTGCAAGCTGAGAGCATCCACCTGCAAG 5334  
Qy 2158 ATACAGACCTGATGTTTGAAGATGACAGATTAACCTCTGTGTGACAGACCTTGAAGA 2217  
Db 5335 ATACAGACCTGATGTTTGAAGATGACAGATTAACCTCTGTGTGACAGACCTTGAAGA 5394  
Qy 2218 ATGCTCATGCGCAACCGTAACTTAAGTCCCTCAACTTTGGAGGACCCACCTGAAGAA 2277  
Db 5395 ATGCTCATGCGCAACCGTAACTTAAGTCCCTCAACTTTGGAGGACCCACCTGAAGAA 5454  
Qy 2278 GAGGATGAAGAGTGGGTGGAAGCTTAAACCCCAAAATGTTGTGAGAGCTTGG 2337  
Db 5455 GAGGATGAAGAGTGGGTGGAAGCTTAAACCCCAAAATGTTGTGAGAGCTTGG 5514  
Qy 2338 AGGCTGATGCTGTGATGATGACCATGCTGTTAAGTGAAGATCTCCCAATCTTAAAG 2397  
Db 5515 AGGCTGATGCTGTGATGATGACCATGCTGTTAAGTGAAGATCTCCCAATCTTAAAG 5574

```

QY 2398 ACCGCCAGCCTGAAATCTCTGAGCCTGGCAGGAAACAAAGTGACAGACCAGGAGTA 2457
    |||
DB 5575 ACCGCCAGCCTGAAATCTCTGAGCCTGGCAGGAAACAAAGTGACAGACCAGGAGTA 5634
    |||
QY 2458 ATGCTCTCAGTGATGAGCCTTGAGAGTCTTCCAGTGGCCCTGACAGAGCTGATCTGAG 2517
    |||
DB 5635 ATGCTCTCAGTGATGAGCCTTGAGAGTCTTCCAGTGGCCCTGACAGAGCTGATCTGAG 5694
    |||
QY 2518 GACTGGGATCAGACGACCGGTTGGCCAGAGTCTGGCCTGAGCCCTGTCAGCAACCGG 2577
    |||
DB 5695 GACTGGGATCAGACGACCGGTTGGCCAGAGTCTGGCCTGAGCCCTGTCAGCAACCGG 5754
    |||
QY 2578 AGCTTGACACACCTTGCTGCTTATCCAAACAACAGCTGGGAAAGAAAGTGAATCTACTG 2637
    |||
DB 5755 AGCTTGACACACCTTGCTGCTTATCCAAACAACAGCTGGGAAAGAAAGTGAATCTACTG 5814
    |||
QY 2638 TGTGCATCCATGAGGCTTCCCACTGTAGTCTGACAGAGGCTGATGCTGAATCAGTCCAC 2697
    |||
DB 5815 TGTGCATCCATGAGGCTTCCCACTGTAGTCTGACAGAGGCTGATGCTGAATCAGTCCAC 5874
    |||
QY 2698 CTGACACGCGCTGAGCTGTGTTTCTTGCACTTGGCCTTATGAGTCAATGGCTGACG 2757
    |||
DB 5875 CTGACACGCGCTGAGCTGTGTTTCTTGCACTTGGCCTTATGAGTCAATGGCTGACG 5934
    |||
QY 2758 CACCTGAGCCTTATGACATGAACCTGTGGAAGCAATGGCGTGAAGCTTCTGTGCGAGGTC 2817
    |||
DB 5935 CACCTGAGCCTTATGACATGAACCTGTGGAAGCAATGGCGTGAAGCTTCTGTGCGAGGTC 5994
    |||
QY 2818 ATGAGAGAACCATCTTGTTCATCTCCAGAGACCTGGAGTTGGTAAAGTGTCTATCCAGCCG 2877
    |||
DB 5995 ATGAGAGAACCATCTTGTTCATCTCCAGAGACCTGGAGTTGGTAAAGTGTCTATCCAGCCG 6054
    |||
QY 2878 GCGTCTGTGAGAGTCTGTCTGCTGTGATCTGAGAGAGACACCTGAAAGAGCTGTGAT 2937
    |||
DB 6055 GCGTCTGTGAGAGTCTGTCTGCTGTGATCTGAGAGAGACACCTGAAAGAGCTGTGAT 6114
    |||
QY 2938 CTGACGACCAATGCTGAGGAGACCGGTGGGTTGCTGCGCTGTGCGAGGACCTGAAGCA 2997
    |||
DB 6115 CTGACGACCAATGCTGAGGAGACCGGTGGGTTGCTGCGCTGTGCGAGGACCTGAAGCA 6174
    |||
QY 2998 AAGAACAGTGTCTGACGAGATCGGGTTGAAGGCAATGTGACCTGATCTGATTTGCTGT 3057
    |||
DB 6175 AAGAACAGTGTCTGACGAGATCGGGTTGAAGGCAATGTGACCTGATCTGATTTGCTGT 6234
    |||
QY 3058 GAGGCACTCTCCTTGAGCCCTTCCGCAACCGGCACTGACCAAGCTAAACCTGGTGGAG 3117
    |||
DB 6235 GAGGCACTCTCCTTGAGCCCTTCCGCAACCGGCACTGACCAAGCTAAACCTGGTGGAG 6294
    |||
QY 3118 AATTAATTCAGTCCCAAGAAATGATGAAGCTGTGGCCCTTGGCTGTCCACGCTCT 3177
    |||
DB 6295 AATTAATTCAGTCCCAAGAAATGATGAAGCTGTGTGGCCCTTGGCTGTCCACGCTCT 6354
    |||
QY 3178 AACTTACAGATTAATGG 3194
    |||
DB 6355 AACTTACAGATTAATGG 6371
    |||

```

## RESULT 10

AA147140  
ID AA147140 standard; DNA; 6939 BP.

AA147140;

20-AUG-2002 (first entry)

Pyrin domain containing protein NALP3/Py17 coding sequence.

XX Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;  
XX antiarteriosclerotic; antipsoriatic; antibacterial; varicoid;  
XX neuroprotective; antiarthritic; antirheumatic; antiashtmatic;  
XX nephrotropic; osteopathic; nootropic; intracellular signal transduction;  
XX inflammation; Alzheimer's disease; infection; psoriasis; asthma;  
XX arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;

```

KW osteoarthritis; glomerulonephritis; gene; ds.
XX Unidentified.
OS
XX WO200240668-A2.
XX
XX PD 23-MAY-2002.
XX
XX PF 30-OCT-2001; 2001WO-BP012545.
XX
XX PR 15-NOV-2000; 2000DE-01056687.
XX
XX PR 30-NOV-2000; 2000DE-01059595.
XX
XX PA (APOT-) APOTECH RES & DEV LTD.
XX
XX PI Teschopp J, Martinson F;
XX
XX WPI; 2002-427093/45.
XX
XX P-PSDB; MA017868.
XX
XX PT New DNA encoding protein with pyrin domain, useful for treating diseases
XX PT involving impaired signal transduction, particularly inflammation, also
XX PT proteins and antibodies.
XX
XX PS Claim 5; Fig 1; 116pp; German.
XX
XX CC The present invention relates the DNA and their encoded proteins, where
XX CC the proteins contain at least one PYD (pyrin) domain. These can be used
XX CC to treat diseases associated with impaired intracellular signal
XX CC transduction, particularly inflammation such as psoriasis,
XX CC arteriosclerosis, bacterial or viral infections (particularly meningitis
XX CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
XX CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
XX CC the invention
XX
XX SQ Sequence 6939 BP; 1791 A; 1719 C; 1801 G; 1628 T; 0 U; 0 Other;
XX
XX
XX Query Match 75.8%; Score 3057.8; DB 6; Length 6939;
XX Best Local Similarity 96.2%; Pred. No. 0;
XX Matches 3192; Conservative 0; Mismatches 2; Indels 123; Gaps 1;
XX
QY 1 ATGGAAGAGACCAATGCTCACCTTTCCAGTACGCGGCTGCATGCTGTCTATGAG 60
    |||
DB 3055 ATGGAAGAGACCAATGCTCACCTTTCCAGTACGCGGCTGCATGCTGTCTATGAG 3114
    |||
QY 61 CTGACCAAGAAAGATTTCAAGACATTTCAAGAAATTTCTTAAGAAAGAAATCTTCGAATCG 120
    |||
DB 3115 CTGACCAAGAAAGATTTCAAGACATTTCAAGAAATTTCTTAAGAAAGAAATCTTCGAATCG 3174
    |||
QY 121 ACCAGATGCTCTATTTCCACAGTTTGAATGCAAGATGCAACGTAAGTGTGGCACTC 180
    |||
DB 3175 ACCAGATGCTCTATTTCCACAGTTTGAATGCAAGATGCAACGTAAGTGTGGCACTC 3234
    |||
QY 181 CTCTTGATGATGATTTATGAGATGCTGGCTGGCTGACCTGCTATGATTCATCTTTGAA 240
    |||
DB 3235 CTCTTGATGATGATTTATGAGATGCTGGCTGGCTGACCTGCTATGATTCATCTTTGAA 3294
    |||
QY 241 AACATGAACCTGCAACCTCTCTGGAGAAAGCAGGAGATGATGAA----- 289
    |||
DB 3295 AACATGAACCTGCAACCTCTCTGGAGAAAGCAGGAGATGATGAAAGAAAGATTCACCA 3354
    |||
QY 290 ----- 289
    |||
DB 3355 GAAGATCCTGAAGCAAGATGACTGACCAAGACCAAGCAAGAAAGTCCAGGAATT 3414
    |||
QY 290 -----AAATTTCA 297
    |||
DB 3415 TCACAAGCTGTGCAACAAGATGATGCAAGCTGCAAGACCAAGAAAGAAATTTCA 3474
    |||
QY 298 CAAGCTATGAAACAAGAGGTGCCACAGCAGCAGAGACAGAGAAACAAGAAATTTCCAA 357
    |||
DB 3475 CAAGCTATGAAACAAGAGGTGCCACAGCAGCAGAGACAGAGAAACAAGAAATTTCCAA 3534
    |||

```

358 GCTATGAAACAAGAGTGCACAGCAGAGAGACAGAAACAAGCATGAGGTGAC 417  
3535 GCTATGAAACAAGAGTGCACAGCAGAGAGACAGAAACAAGCATGAGGTGAC 3594  
418 ACATGGAGCTACAGAGTACAGTATGACCAAAATTCGTGAGAGAGAGATGACCTGCT 477  
3595 ACATGGAGCTACAGAGTACAGTATGACCAAAATTCGTGAGAGAGAGATGACCTGCT 3654  
478 AGTTTGAACACATGCTGCTGACCTGCGGAAATTCGAAACGTTGCTGCTGCTTTGAT 537  
3655 AGTTTGAACACATGCTGCTGACCTGCGGAAATTCGAAACGTTGCTGCTGCTTTGAT 3714  
538 TCAGACCGGTGGGGCTTCCGGCTCGCAGAGGTGTTCTGACGGAAGTCAGGAATTTGGG 597  
3715 TCAGACCGGTGGGGCTTCCGGCTCGCAGAGGTGTTCTGACGGAAGTCAGGAATTTGGG 3774  
598 AATCGGCTCTAGCAGAGAGATCGTGTGCTGGCGCAAGTGAAGCTTACAGAGGA 657  
3775 AATCGGCTCTAGCAGAGAGATCGTGTGCTGGCGCAAGTGAAGCTTACAGAGGA 3834  
658 AGTTCTCTACGCTCTTCTTCTCCCGTTAGAGATGACAGCGGAAGAGAGACAGT 717  
3835 AGTTCTCTACGCTCTTCTTCTCCCGTTAGAGATGACAGCGGAAGAGAGACAGT 3894  
718 GTACAGAGTTCATCTCCAGAGAGTGGCCAGACTCCAGAGCTCCGGTACGAGATCAG 777  
3895 GTACAGAGTTCATCTCCAGAGAGTGGCCAGACTCCAGAGCTCCGGTACGAGATCAG 3954  
778 TCCCGACAGAAAGGCTGTTGTTTCATCATTTGACGTTGATGACCTGGGCTCTGCTC 837  
3955 TCCCGACAGAAAGGCTGTTGTTTCATCATTTGACGTTGATGACCTGGGCTCTGCTC 4014  
838 AACATGACACAAAGCTCTGCAAAAGCTGGGCTGAGAGACAGCTCCGTTACCTTCATA 897  
4015 AACATGACACAAAGCTCTGCAAAAGCTGGGCTGAGAGACAGCTCCGTTACCTTCATA 4074  
898 CGCAGTCTGCTAGAGAGTCTGCTCCCTGAGTCTTCTGATGTCACGTCAGAGAC 957  
4075 CGCAGTCTGCTAGAGAGTCTGCTCCCTGAGTCTTCTGATGTCACGTCAGAGAC 4134  
958 GTGGGACAGAGAGTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTTAGTTAGAGA 1017  
4135 GTGGGACAGAGAGTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTTAGTTAGAGA 4194  
1018 ATCTCCGGGAAACAAGATTCACCTTGTCTTGAACGCGGAGTTGGTAGATCAGAAG 1077  
4195 ATCTCCGGGAAACAAGATTCACCTTGTCTTGAACGCGGAGTTGGTAGATCAGAAG 4254  
1078 ACACAAGGTTGCGTGCATCATGAAACAACGCTGAGCTGTCGACAGTGCAGGTCGCC 1137  
4255 ACACAAGGTTGCGTGCATCATGAAACAACGCTGAGCTGTCGACAGTGCAGGTCGCC 4314  
1138 GCGGTGGGCTCTCATCTGCGTGGCCGTCGAGCTGACGAGAGTGGTGGGGAGAGCTC 1197  
4315 GCGGTGGGCTCTCATCTGCGTGGCCGTCGAGCTGACGAGAGTGGTGGGGAGAGCTC 4374  
1198 GCCCCTTCAACCAACGCTCAACAGGCTGACAGCGCTTTTGTGTTTCATCAGCTCAC 1257  
4375 GCCCCTTCAACCAACGCTCAACAGGCTGACAGCGCTTTTGTGTTTCATCAGCTCAC 4434  
1258 CCTGAGGCGTGGTCCGCGCTGTCTCAATCTGAGAGAAAGTTGCTCTGAAGCGCTTC 1317  
4435 CCTGAGGCGTGGTCCGCGCTGTCTCAATCTGAGAGAAAGTTGCTCTGAAGCGCTTC 4494  
1318 TGCCGATGAGCTGTGAGAGGAGTGAATAGGAAGTCAGTGTGAGTGAAGAGCTC 1377  
4495 TGCCGATGAGCTGTGAGAGGAGTGAATAGGAAGTCAGTGTGAGTGAAGAGCTC 4554  
1378 ATGTTTCAAGAGCTCGGGAGTCTGAGCTCCGTGCTGTGTTTCAATGAAACATCTTCTC 1437  
4555 ATGTTTCAAGAGCTCGGGAGTCTGAGCTCCGTGCTGTGTTTCAATGAAACATCTTCTC 4614

1438 CCAGACAGCTAGTGAAGAGTACTACACTTCTTCCAGCTCAGTCTCCAGAGCTTCTGT 1497  
4615 CCAGACAGCTAGTGAAGAGTACTACACTTCTTCCAGCTCAGTCTCCAGAGCTTCTGT 4674  
1498 GCCGCTTGTACTAGTGTGAAGGCTGGAATTCAGGACGCTCTGTGCTCTGTAC 1557  
4675 GCCGCTTGTACTAGTGTGAAGGCTGGAATTCAGGACGCTCTGTGCTCTGTAC 4734  
1558 GTTGAAGACAAAGAGTCCATGAGAGCTTAAACAGGAGCTTCCATATCACTGCTT 1617  
4735 GTTGAAGACAAAGAGTCCATGAGAGCTTAAACAGGAGCTTCCATATCACTGCTT 4794  
1618 TGATGAAGCTTCTTGTGAGCTCGTGAAGCAAGTAAAGAGGCTCAGTGAAGGTC 1677  
4795 TGATGAAGCTTCTTGTGAGCTCGTGAAGCAAGTAAAGAGGCTCAGTGAAGGTC 4854  
1678 CTGCTGGGCTGTCCCGTTCCTCGGGGTGAAGCAAGAGCTTTCAGTGGGTCTCTGT 1737  
4855 CTGCTGGGCTGTCCCGTTCCTCGGGGTGAAGCAAGAGCTTTCAGTGGGTCTCTGT 4914  
1738 TTGGGTACAGAGCTTAATGCCACACCCAGAGACACCTTGAAGCTTCCACTGTCTT 1797  
4915 TTGGGTACAGAGCTTAATGCCACACCCAGAGACACCTTGAAGCTTCCACTGTCTT 4974  
1798 TTGAGACTCAAGACAAAGAGTTGTTGCTTGGCATTAACAGCTTCCAAAGAGTGG 1857  
4975 TTGAGACTCAAGACAAAGAGTTGTTGCTTGGCATTAACAGCTTCCAAAGAGTGG 5034  
1858 CTTCCGATTAACAGAACTGGAAGCTTGAATGATCTTCTGCTCCAGACAGTGTCCG 1917  
5035 CTTCCGATTAACAGAACTGGAAGCTTGAATGATCTTCTGCTCCAGACAGTGTCCG 5094  
1918 TATTTGCGAAATTCGGGTGATGTCAAAGGATCTTCCAGAGATGAGTCCGCTGAG 1977  
5095 TATTTGCGAAATTCGGGTGATGTCAAAGGATCTTCCAGAGATGAGTCCGCTGAG 5154  
1978 GCATGCTCTGTGTCCTCTATGATGATGGGATTAAGACCTCATTTAGAGACAGTGGAA 2037  
5155 GCATGCTCTGTGTCCTCTATGATGATGGGATTAAGACCTCATTTAGAGACAGTGGAA 5214  
2038 GATTTCTGCTCATGCTTGGACCAACCACTGCGGAGCTGAGCCTGGGAGCAGC 2097  
5215 GATTTCTGCTCATGCTTGGACCAACCACTGCGGAGCTGAGCCTGGGAGCAGC 5274  
2098 ATCTGACAGAGCGGACATGAAGACCTGTGTGCAAGCTGAGAGATCCACCTGCAAG 2157  
5275 ATCTGACAGAGCGGACATGAAGACCTGTGTGCAAGCTGAGAGATCCACCTGCAAG 5334  
2158 ATACAGACCTGATGTTTGAAGATGACAGATTAACCTGCTGTGTGACAGACCTTGGAGA 2217  
5335 ATACAGACCTGATGTTTGAAGATGACAGATTAACCTGCTGTGTGACAGACCTTGGAGA 5394  
2218 ATGCTATGAGCAACCGTAACTTAAGATCCCTCAACTTGGAGGACCACTGAAGAA 2277  
5395 ATGCTATGAGCAACCGTAACTTAAGATCCCTCAACTTGGAGGACCACTGAAGAA 5454  
2278 GAGATGTAAGATGCGCTGTAAGCTTAAACACCCAAATGTTTGTGAGTCTTTG 2337  
5455 GAGATGTAAGATGCGCTGTAAGCTTAAACACCCAAATGTTTGTGAGTCTTTG 5514  
2338 AGGCTGATGCTGTGATGTAACCAATGCTGTTAAGTAAAGATCTCCAAATCTTTAG 2397  
5515 AGGCTGATGCTGTGATGTAACCAATGCTGTTAAGTAAAGATCTCCAAATCTTTAG 5574  
2398 ACTTCCCGACCTGAATCTGAGAGCTGACAGAAACAAGTGAAGACAGAGAGTA 2457  
5575 ACTTCCCGACCTGAATCTGAGAGCTGACAGAAACAAGTGAAGACAGAGAGTA 5634  
2458 ATGCTCTCAGTGAATGCTTGAAGAGTCTCCAGTGCAGCTTGAAGAGCTGATACTGAG 2517  
5635 ATGCTCTCAGTGAATGCTTGAAGAGTCTCCAGTGCAGCTTGAAGAGCTGATACTGAG 5694  
2518 GACTGTGATCAAGCAGCAGGAGTTGCCAAGTGTGAGCTGAGCTTGTGACGACCGG 2577

Db 5695 GACCTGGCATCAGACCAACGGGTGGCCAGAGTCTGGCCCTCAGCCCTGTCAGCAACGG 5754  
Qy 2578 AGCTTGACACACCTGTGCTATTCACAAACAAGCCGGGGAAGAGTGTAAATCTACTG 2637  
Db 5755 AGCTTGACACACCTGTGCTATTCACAAACAAGCCGGGGAAGAGTGTAAATCTACTG 5814  
Qy 2638 TGTGCATCATAGAGGCTTCCCACTGTAGTCTGACAGAGGCTGATCTGAATCAGTCCAC 2697  
Db 5815 TGTGCATCATAGAGGCTTCCCACTGTAGTCTGACAGAGGCTGATCTGAATCAGTCCAC 5874  
Qy 2698 CTGACACGGCTGCGCTGTGTCTTTCTTGACCTTGGCTTAAAGTGTATCTCATGCTGACG 2757  
Db 5875 CTGACACGGCTGCGCTGTGTCTTTCTTGACCTTGGCTTAAAGTGTATCTCATGCTGACG 5934  
Qy 2758 CACCTGACGCTTGAATGAACCTGTGAAAGCAATGACGTGAACCTTCTGTGCGAGGTC 2817  
Db 5935 CACCTGACGCTTGAATGAACCTGTGAAAGCAATGACGTGAACCTTCTGTGCGAGGTC 5994  
Qy 2818 ATGAGAGAACCATCTTGTATCTCCAGGACCTGAGTGTAAAGTGTATCTCATGCTGACG 2877  
Db 5995 ATGAGAGAACCATCTTGTATCTCCAGGACCTGAGTGTAAAGTGTATCTCATGCTGACG 6054  
Qy 2878 GCGTCTGTGAGAGTCTGTCTGTGTGATCTTCAGAGACAGACACCTGAAGAGCTTGAT 2937  
Db 6055 GCGTCTGTGAGAGTCTGTCTGTGTGATCTTCAGAGAGACAGACACCTGAAGAGCTTGAT 6114  
Qy 2938 CTGACGGAACAATGCCCTGGGTGACGGTGGGTGTCTGGCTGTGAGAGGACTGAAGAA 2997  
Db 6115 CTGACGGAACAATGCCCTGGGTGACGGTGGGTGTCTGGCTGTGAGAGGACTGAAGAA 6174  
Qy 2998 AAGAACAGTGTCTGACGAGACTCGGGTGAAGGATGGAATGACTCTGATTTGCTGT 3057  
Db 6175 AAGAACAGTGTCTGACGAGACTCGGGTGAAGGATGGAATGACTCTGATTTGCTGT 6234  
Qy 3058 GAGGACCTCTCTTGGCCCTTCTGCAACGGGATCTGACCAAGTCTAAACCTGTGCGAG 3117  
Db 6235 GAGGACCTCTCTTGGCCCTTCTGCAACGGGATCTGACCAAGTCTAAACCTGTGCGAG 6294  
Qy 3118 AATACTTCAGTCCCAAGGAATGATGAAGCTGTGCGCTTGGCTTGGCCCAAGCTCT 3177  
Db 6295 AATACTTCAGTCCCAAGGAATGATGAAGCTGTGCGCTTGGCTTGGCCCAAGCTCT 6354  
Qy 3178 AACTTACAGATTAATGG 3194  
Db 6355 AACTTACAGATTAATGG 6371

RESULT 11  
ABX97181  
ID ABX97181 standard; cDNA; 3226 BP.  
XX  
AC ABX97181;  
XX  
DT 20-MAY-2003 (first entry)  
XX  
DE Human NOV125b cDNA.  
XX  
KM NOXA; cytosolic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;  
KM hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;  
KM human; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200272757-A2.  
XX  
PD 19-SEP-2002.  
XX  
PF 08-MAR-2002; 2002MO-US006908.  
XX  
PR 08-MAR-2001; 2001US-0274101P.  
PR 08-MAR-2001; 2001US-0274194P.  
PR 08-MAR-2001; 2001US-0274281P.

PR 08-MAR-2001; 2001US-0274322P.  
PR 09-MAR-2001; 2001US-0274849P.  
PR 12-MAR-2001; 2001US-0275235P.  
PR 13-MAR-2001; 2001US-0275578P.  
PR 13-MAR-2001; 2001US-0275579P.  
PR 13-MAR-2001; 2001US-0275601P.  
PR 14-MAR-2001; 2001US-0276000P.  
PR 16-MAR-2001; 2001US-0276776P.  
PR 19-MAR-2001; 2001US-0276994P.  
PR 20-MAR-2001; 2001US-0277239P.  
PR 20-MAR-2001; 2001US-0277321P.  
PR 20-MAR-2001; 2001US-0277327P.  
PR 21-MAR-2001; 2001US-0277791P.  
PR 22-MAR-2001; 2001US-0277833P.  
PR 23-MAR-2001; 2001US-0278152P.  
PR 26-MAR-2001; 2001US-0278894P.  
PR 27-MAR-2001; 2001US-0278999P.  
PR 27-MAR-2001; 2001US-0279036P.  
PR 28-MAR-2001; 2001US-0279344P.  
PR 30-MAR-2001; 2001US-0277338P.  
PR 30-MAR-2001; 2001US-0279995P.  
PR 30-MAR-2001; 2001US-0280233P.  
PR 02-APR-2001; 2001US-0280802P.  
PR 02-APR-2001; 2001US-0280822P.  
PR 02-APR-2001; 2001US-0280900P.  
PR 04-APR-2001; 2001US-0281194P.  
PR 13-APR-2001; 2001US-0283675P.  
PR 30-APR-2001; 2001US-0287424P.  
PR 02-MAY-2001; 2001US-0288066P.  
PR 03-MAY-2001; 2001US-0288342P.  
PR 03-MAY-2001; 2001US-0288528P.  
PR 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291099P.  
PR 16-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
PR 31-MAY-2001; 2001US-0294889P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 18-JUN-2001; 2001US-0299027P.  
PR 19-JUN-2001; 2001US-0299303P.  
PR 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
PR 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
PR 10-SEP-2001; 2001US-0318462P.  
PR 12-SEP-2001; 2001US-0318770P.  
PR 27-SEP-2001; 2001US-0325430P.  
PR 27-SEP-2001; 2001US-0325681P.  
PR 18-OCT-2001; 2001US-0330380P.  
PR 31-OCT-2001; 2001US-0335301P.  
PR 14-NOV-2001; 2001US-0332172P.  
PR 14-NOV-2001; 2001US-0332271P.  
PR 14-NOV-2001; 2001US-0332272P.  
PR 14-NOV-2001; 2001US-033184P.  
PR 14-NOV-2001; 2001US-033272P.  
PR 21-NOV-2001; 2001US-0332094P.  
PR 03-DEC-2001; 2001US-0337426P.  
PR 03-DEC-2001; 2001US-0338092P.  
PR 04-DEC-2001; 2001US-0337185P.  
PR 03-JAN-2002; 2002US-0345705P.  
PR 07-MAR-2002; 2002US-00092900.  
  
(CURA-) CURAGEN CORP.  
XX  
PA Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L;  
XX Zethuusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;  
PI Paturajan M, Gangoli E, Vernet CM, Guo X, Tchernev V;  
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;  
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsbrook JP;  
PI Lepley DM, Rieger DK;  
DR WPI; 2002-72332/78.  
DR P-PSDB; ABU65214.  
XX





Db 2049 CACCTCTGGAGATCGTCATGGCCAAACGTAACCTAATCCCTCACTGGGAGGACCC 2108  
Qy 2266 CACCTGAAGAAAGAGATGTAAAGATGGCGTGTGAAGCCTTAAACACCCAAAAGTTTG 2325  
Db 2109 CACCTGAAGAAAGAGATGTAAAGATGGCGTGTGAAGCCTTAAACACCCAAAAGTTTG 2168  
Qy 2326 TTGAGTCTTTGAGGCTGGAATTGCTGTGGAATTGACCCATGCGCTGTACTGTAAAGATCTCC 2385  
Db 2169 TTGAGTCTTTGAGGCTGGAATTGCTGTGGAATTGACCCATGCGCTGTACTGTAAAGATCTCC 2228  
Qy 2386 CAAATCTTACGACCTCCCCAGCCTGAATCTGTAGCCTTGGCAGAAACAAGGTGACA 2445  
Db 2229 CAAATCTTACGACCTCCCCAGCCTGAATCTGTAGCCTTGGCAGAAACAAGGTGACA 2288  
Qy 2446 GACCGAGGAGTAATGCTCTGATGATGCTTGAAGTCTCCCAATGGCGCCTGAGAGAG 2505  
Db 2289 GACCGAGGAGTAATGCTCTGATGATGCTTGAAGTCTCCCAATGGCGCCTGAGAGAG 2348  
Qy 2506 CTGATCTGAGAGACTGTGACATCAACAGCGGTTCCAGAGTCTGGCCTCAGCCCTC 2565  
Db 2349 CTGATCTGAGAGACTGTGACATCAACAGCGGTTCCAGAGTCTGGCCTCAGCCCTC 2408  
Qy 2566 GTGAGCAACCGAGCTTGAACAACCTGTGCTATCAACAACAGCTGAGGAAACGAGGT 2625  
Db 2409 GTGAGCAACCGAGCTTGAACAACCTGTGCTATCAACAACAGCTGAGGAAACGAGGT 2468  
Qy 2626 GTAAATCTAATGTCATGATCCATGAGGCTTCCCCCTGTAATGTCAGAGGCTGATGCTG 2685  
Db 2469 GTAAATCTAATGTCATGATCCATGAGGCTTCCCCCTGTAATGTCAGAGGCTGATGCTG 2528  
Qy 2686 AATCAGTCCACTGAGACAGCGCTGTGAGTTTCTTGTGCACTTGCCCTTATGAGGTAC 2745  
Db 2529 AATCAGTCCACTGAGACAGCGCTGTGAGTTTCTTGTGCACTTGCCCTTATGAGGTAC 2588  
Qy 2746 TCATGCTGAGCGACCTGAGCCTTAGCATGAACCTCTGGAAGACATGCGGTGAAGCTT 2805  
Db 2589 TCATGCTGAGCGACCTGAGCCTTAGCATGAACCTCTGGAAGACATGCGGTGAAGCTT 2648  
Qy 2806 CTGTGCGAGGTCATGAGAAACATCTTGTGATCTCCGAGACCTGAGATGTTGTAAGGT 2865  
Db 2649 CTGTGCGAGGTCATGAGAAACATCTTGTGATCTCCGAGACCTGAGATGTTGTAAGGT 2708  
Qy 2866 CATCTCACCGCGCGCTGTGAGAGTCTGTCTGTGATCTCGAGAGACACCTG 2925  
Db 2709 CATCTCACCGCGCGCTGTGAGAGTCTGTCTGTGATCTCGAGAGACACCTG 2768  
Qy 2926 AAGAGCTGTGATCTCAACGACATGCTCGGTGACGCTGGGTTGCTGCGCTGTGCGAG 2985  
Db 2769 AAGAGCTGTGATCTCAACGACATGCTCGGTGACGCTGGGTTGCTGCGAG 2828  
Qy 2986 GGAAGTGAAGAAAGACAGTGTCTGAAGACCTCGGTTGAAGGATGTGAGACTGACT 3045  
Db 2829 GGAAGTGAAGAAAGACAGTGTCTGAAGACCTCGGTTGAAGGATGTGAGACTGACT 2888  
Qy 3046 TCTGATGCTGTGAGGCACTCTCTTGGCCCTTCTCTCAACCGGCACTGACAGTCTA 3105  
Db 2889 TCTGATGCTGTGAGGCACTCTCTTGGCCCTTCTCTCAACCGGCACTGACAGTCTA 2948  
Qy 3106 AACCTGTGCAAGATTAATCTTCAATGTCCTCAAGAGATGTAAGCTGTGGCCTTTGCC 3165  
Db 2949 AACCTGTGCAAGATTAATCTTCAATGTCCTCAAGAGATGTAAGCTGTGGCCTTTGCC 3008  
Qy 3166 TGTCCAGCTCTAATTAATTAAGATTAATGAGGCTGTGGAATGACAGTACCTGTGCAATA 3225  
Db 3009 TGTCCAGCTCTAATTAATTAAGATTAATGAGGCTGTGGAATGACAGTACCTGTGCAATA 3068  
Qy 3226 AGAAGTGTGAGAGAGTGAAGTGAAGTCAAGCCCGAGTGTGAATGAAGGTAGTTGG 3285  
Db 3069 AGAAGTGTGAGAGAGTGAAGTGAAGTCAAGCCCGAGTGTGAATGAAGGTAGTTGG 3128  
Qy 3286 CATCTTTTGAATGAAGATGACCGACAC 3312  
PR |||||

Db 3129 CATCTTTGATGAAGATGACCGGTAC 3155  
RESULT 12  
ID ADN62078 standard; CDNA; 2820 BP.  
XX  
AC ADN62078;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Human CDNA encoding NOV125a.  
XX  
KW Human; ss; gene; NOXV; diabetes; obesity; infectious disease; anorexia;  
KW cancer-associated cachexia; cancer; neurodegenerative disorder;  
KW Alzheimer's disease; Parkinson's disease; immune disorder;  
KW haematopoietic disorder; dyslipidaemia; chronic disease.  
XX  
OS Homo sapiens.  
XX  
PN US2004043382-A1.  
XX  
PD 04-MAR-2004.  
XX  
PF 07-MAR-2002; 2002US-00092900.  
XX  
PR 08-MAR-2001; 2001US-0274191P.  
PR 08-MAR-2001; 2001US-0274194P.  
PR 08-MAR-2001; 2001US-0274281P.  
PR 08-MAR-2001; 2001US-0274322P.  
PR 09-MAR-2001; 2001US-0274849P.  
PR 12-MAR-2001; 2001US-0275235P.  
PR 13-MAR-2001; 2001US-0275578P.  
PR 13-MAR-2001; 2001US-0275579P.  
PR 13-MAR-2001; 2001US-0275601P.  
PR 14-MAR-2001; 2001US-0276000P.  
PR 16-MAR-2001; 2001US-0276776P.  
PR 19-MAR-2001; 2001US-0276994P.  
PR 20-MAR-2001; 2001US-0277239P.  
PR 20-MAR-2001; 2001US-0277321P.  
PR 20-MAR-2001; 2001US-0277327P.  
PR 20-MAR-2001; 2001US-0277338P.  
PR 21-MAR-2001; 2001US-0277791P.  
PR 22-MAR-2001; 2001US-0277833P.  
PR 23-MAR-2001; 2001US-0278152P.  
PR 26-MAR-2001; 2001US-0278894P.  
PR 27-MAR-2001; 2001US-0278999P.  
PR 27-MAR-2001; 2001US-0279036P.  
PR 28-MAR-2001; 2001US-0279344P.  
PR 30-MAR-2001; 2001US-0279959P.  
PR 30-MAR-2001; 2001US-0280233P.  
PR 02-APR-2001; 2001US-0280802P.  
PR 02-APR-2001; 2001US-0280822P.  
PR 02-APR-2001; 2001US-0280900P.  
PR 04-APR-2001; 2001US-0281444P.  
PR 13-APR-2001; 2001US-0283675P.  
PR 30-APR-2001; 2001US-0287424P.  
PR 02-MAY-2001; 2001US-0288066P.  
PR 03-MAY-2001; 2001US-0288342P.  
PR 03-MAY-2001; 2001US-0288528P.  
PR 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291099P.  
PR 16-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
PR 31-MAY-2001; 2001US-0294889P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 18-JUN-2001; 2001US-0299027P.  
PR 19-JUN-2001; 2001US-0299303P.  
PR 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
PR 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
PR 10-SEP-2001; 2001US-0318462P.



853 TCCAGCTGCCGCCG----- 868  
868  
1186 GGGGAGAGGTGCCCCCTTCAACCAAGCTCAGAGGCTGACGCGCTTTTGTGTTT 1245  
869 ----- 868  
1246 CATCAGCTCACCCCTCGAGGCGTGTCCGCGCTGTCTCAATCTGGAGAAAAGTGTGTC 1305  
869 ----- 868  
1306 CTGAAGCCCTTCTCCGCTATGCTGTGAGGGAGTGTGAATAGGAAGTCAGTGTTCAT 1365  
869 -----TCTGCCGATGCTGTGAGGGAGTGTGAATAGGAAGTCAGTGTTCATC 918  
1366 GGTGACGACTCATGATGATCAAGGACTCGGGGAGTCTGAGCTCCGATCTGTGTTTCAATG 1425  
919 GGTGACGACTCATGATGATCAAGGACTCGCGAGTCTGAGCTCCGATCTGTGTTTCAATG 978  
1426 AACATCCTTCTCCAGACAGCCACTGTGAGAGTACTACACCTTCTTCCACTCAGTCTC 1485  
979 AACATCCTTCTCCAGACAGCCACTGTGAGAGTACTACACCTTCTTCCACTCAGTCTC 1038  
1486 CAGGACTTCTGTGCGCCCTTGTACTACGCTTTAGAGGGCTGTGAATATGAGCAGCTCTC 1545  
1039 CAGGACTTCTGTGCGCCCTTGTACTACGCTTTAGAGGGCTGTGAATATGAGCAGCTCTC 1098  
1546 TGCCCTCTGTAGCTTGAAGAGCAAAAGAGTCAATGAGGCTTAAACAGGAGGCTTCAT 1605  
1099 TGCCCTCTGTAGCTTGAAGAGCAAAAGAGTCAATGAGGCTTAAACAGGAGGCTTCAT 1158  
1606 ATCCACTGCTTTGGATGAAGGCTTCTGTTTGGCTCTGTGAGGAAAGCTTAAGAGAG 1665  
1159 ATCCACTGCTTTGGATGAAGGCTTCTGTTTGGCTCTGTGAGGAAAGCTTAAGAGAG 1218  
1666 CCACTGAGAGTCTCTGTGAGGCTGTGCTCCGCTTCCCTGAGGGGTGAAGCAAGCTTTCAC 1725  
1219 CCACTGAGAGTCTCTGTGAGGCTGTGCTCCGCTTCCCTGAGGGGTGAAGCAAGCTTTCAC 1278  
1726 TGGGCTCTGCTTTGGGTCAGAGGCTTAATGCAACCCAGAGAGACCCCTGAGAGCC 1785  
1279 TGGGCTCTGCTTTGGGTCAGAGGCTTAATGCAACCCAGAGAGACCCCTGAGAGCC 1338  
1786 TTCCACTGCTTTTGGAGACTCAAGACAAAGTGTTCGCTTGGCAATTAAACAGTTC 1845  
1339 TTCCACTGCTTTTGGAGACTCAAGACAAAGTGTTCGCTTGGCAATTAAACAGTTC 1398  
1846 CAAGAAGTGTGCTTCCGATTAAACGAACCTTGAATGATCTTCTCTGCTC 1905  
1339 CAAGAAGTGTGCTTCCGATTAAACGAACCTTGAATGATCTTCTCTGCTC 1458  
1906 CAGCACTGCTGATTTTGGGGAATTTGGGGTGAATGCAAAAGGATCTTCCCAAGAT 1965  
1459 CAGCACTGCTGATTTTGGGGAATTTGGGGTGAATGCAAAAGGATCTTCCCAAGAT 1518  
1966 GAGTCCGCTGAGGACTGTCTGTGTCTCTCTATGATGCGGATTAAGACCTCAATTAG 2025  
1519 GAGTCCGCTGAGGACTGTCTGTGTCTCTCTATGATGCGGATTAAGACCTCAATTAG 1578  
2026 GAGCACTGAGGAATTTTGTCTCTCAATGCTTGGCAACCCACCACTGCGGAGCTGAC 2085  
1579 GAGCACTGAGGAATTTTGTCTCTCAATGCTTGGCAACCCACCACTGCGGAGCTGAC 1624  
2086 CTGGGACAGAGATCTGTGACAGAGGCGGCAATGAAGACCTGTGTGCGCAAGCTGAGGAT 2145  
1625 -----TGCAAGCTGAGGAT 1640  
2146 CCCACTGCAAGATACAGACCTGTATGTTTAAAGATGACAGATTAACCTGTGTGTGAG 2205  
1641 CCCACTGCAAGATACAGACCTGTATGTTTAAAGATGACAGATTAACCTGTGTGTGAG 1700  
2206 CACTCTGAGAAATGTGATGAGGCAACCGTAACCTTAAGATCTTCAACTTTGGAGGAC 2265  
1701 CACTCTGAGAAATGTGATGAGGCAACCGTAACCTTAAGATCTTCAACTTTGGAGGAC 1760

2266 CACTGAGGAAGAGATGTAAGATGAGCGGTGAAGCCTTAAACACCCAAATGTTT 2325  
1761 CACTCAACGAAGAGATGTAAGATGAGCGGTGAAGCCTTAAACACCCAAATGTTT 1820  
2326 TTGAGTCTTTGAGGCTGAGATTTGCTGTGATTTGAACCGATGCTTTAAGATCTCC 2385  
1821 TTGAGTCTTTGAGGCTGAGATTTGCTGTGATTTGAACCGATGCTTTAAGATCTCC 1880  
2386 CAAATCTTACGACTTCCCGAGCTGAATTTCTGAGCCTGAGAGAAACAAAGTACA 2445  
1881 CAAATCTTACGACTTCCCGAGCTGAATTTCTGAGCCTGAGAGAAACAAAGTACA 1940  
2446 GACCAAGGAGTAATGCTCTCAGATGATGCTTGAAGTCTCCAGTGCAGCCTGACAG 2505  
1941 GACCAAGGAGTAATGCTCTCAGATGATGCTTGAAGTCTCCAGTGCAGCCTGACAG 2000  
2506 CTGATCTGAGAGACTGTGATCAACGCAAGGCTTGAAGTCTGAGCTCAGCCCTC 2565  
2001 CTGATCTGAGAGACTGTGATCAACGCAAGGCTTGAAGTCTGAGCTCAGCCCTC 2060  
2566 GTGAGCAACCGGAGCTTGAACACCTGTGCTTATCCAAACAGCCTGGGAAAGAGT 2625  
2061 GTGAGCAACCGGAGCTTGAACACCTGTGCTTATCCAAACAGCCTGGGAAAGAGT 2120  
2626 GTAAATCTACTGTGTGATCATGAGGCTTCCCACTGTAGTCTGACAGGCTGATCTG 2685  
2121 GTAAATCTACTGTGTGATCATGAGGCTTCCCACTGTAGTCTGACAGGCTGATCTG 2180  
2686 AATCAGTGCACCTGGAACAGGCTGTGCTGCTTTTCTTGAACCTTGAAGGCTTAAG 2745  
2181 AATCAGTGCACCTGGAACAGGCTGTGCTGCTTTTCTTGAACCTTGAAGGCTTAAG 2240  
2746 TCAATGCTGAGCACTGAGCCTTGAAGATGAACCTTGAAGCAATGAGGTGAAGT 2805  
2241 TCAATGCTGAGCACTGAGCCTTGAAGATGAACCTTGAAGCAATGAGGTGAAGT 2300  
2806 CTGTGAGAGTCAATGAGAAACATCTTGTATCTCCAGACCTGAGGCTTGAAGT 2865  
2301 CTGTGAGAGTCAATGAGAAACATCTTGTATCTCCAGACCTGAGGCTTGAAGT 2360  
2866 CATCTCACCGCGGTGTGAGAGTCTGTCTGTGTGATCTTGAAGAGACACCTG 2925  
2361 CATCTCACCGCGGTGTGAGAGTCTGTCTGTGTGATCTTGAAGAGACACCTG 2378  
2926 AAGAGCTGATCTCAGGGAACAATGCCCTGTGAGACGCTGGGCTTGCCTGTGTGAG 2985  
2379 -----CGAACAATGCCCTGTGAGACGCTGGGCTTGCCTGTGTGAG 2422  
2986 GGACTGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGATGAGTACT 3045  
2423 GGACTGAAGCAAAAGAACAGTGTCTGACGAGACTCGGGTTGAAGGATGAGTACT 2482  
3046 TGTATGCTGTGAGGACTCTCTCTGAGCCTTCTGCAACCGGACTTGAACGATCTA 3105  
2483 TGTATGCTGTGAGGACTCTCTCTGAGCCTTCTGCAACCGGACTTGAACGATCTA 2542  
3106 AACTGTGAGGATTAATCAATGCTTGAAGGATTAATGAGTGTGAGGCTTGGCC 3165  
2543 AACTGTGAGGATTAATCAATGCTTGAAGGATTAATGAGTGTGAGGCTTGGCC 2602  
3166 TGTCCAGCTTAATTAAGATTAATGAGGCTTGAAGGATTAATGAGGCTTGAAGT 3225  
2603 TGTCCAGCTTAATTAAGATTAATGAGGCTTGAAGGATTAATGAGGCTTGAAGT 2662  
3226 AGGAAGCTGTGAGGAGAGTGAAGTGAAGGCTTGAAGGCTTGAAGGCTTGAAGT 3285  
2663 ACCAAGCTGTGAGGAGAGTGAAGTGAAGGCTTGAAGGCTTGAAGGCTTGAAGT 2722  
3286 CATCTTTTGAAGAGTGAAGTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 3312  
2723 CATCTTTTGAAGAGTGAAGTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 2749



Db 1646 CTATTTGGCCTCATGAACAGAGATATCTTGAGAGCTTCGTTTGATATCC 1705  
 Qy 1692 GCTTCCTCGGGGGGGAAGACAGAGCTTCTGCACTGGGCTCTCTGTTGGGACAGACC 1751  
 Db 1706 CGTGATTCACACTGTGTGAGCAGAGAGCTCCACACCTGGGCTCTCTGATAGCTCAGACAGT 1765  
 Qy 1752 TAATGCAACACCCGAGAGACACCTGAGCGCTTCACTGTCTTTTTCAGACTCAAGA 1811  
 Db 1766 CAATGACACAGCCCAATGAGACACCTGGATGCTCTTATTGTCTAATTTGATGCTCAGGA 1825  
 Qy 1812 CAAGAAGTTGTGCTGCTGCACTTAACAGCTTCCAAAGATGATGGCTTCGATTAACCA 1871  
 Db 1836 TGAAGAAGTTGTGGCGGGGCTCCAAACGCTTCCAAAGATGATGGCTGCTGATTAACCA 1885  
 Qy 1872 GAACCTGAGACTTGAATGAGCTTCTCTGCTGCTCCAGAGCTGTCCGTAATTTGGCGGAAT 1931  
 Db 1886 GAAAGATGACTTGAAGGCTCTCTCTCTACTGTCTCAGACCTGTCAAGAACTTGAAGGCAAT 1945  
 Qy 1932 TCGGGTGAATGCAAAAGGATCTTCCAAAGATGATGGCTGCTGATGCTGCTGCTGCTGCT 1991  
 Db 1946 CCGGGTGAATGCAAGAGCTCTCTCGGTAGATTAATCTCTGAGCTGTGCTGCTGCTGCTGCT 2005  
 Qy 1992 CCTCTATGATGAGGGGA---TAAGACCTCATTTAGAGAGCAGTGGGAAAGATTTGCTGCTC 2048  
 Db 2006 TACTGTCCAGAGACACATATGATGACCCCTCTCTCATGAGTGGTGGGAAACCTTCTGCTC 2065  
 Qy 2049 CATGCTTGGACCCACCCACACCTGCGGAGCTGAGACCTGGGACAGACATCTTGACAGA 2108  
 Db 2066 TGTGCTTGGACCTCTCCGAACTTGAAGAGCTGAGACTTGGGACAGACATCTTGAGTCA 2125  
 Qy 2109 GCGGGCCATGAAGACCTGTGTGCCAAGCTGAGCATCCCATGCAAGATACAGACCT 2168  
 Db 2126 ACGGGCCATGAAGATACTGTGCTGAGCTGCGAATCAGTCTCAGAAATACAGAACT 2185  
 Qy 2169 GATGTTGAAGATGACAGATTAACCTGTGTGTGAGACACTCTGAGAAATGCTATGAC 2228  
 Db 2186 GACGTTGAAGATGAGAGTATGCTGTGCTGCTGAACATCTCTGAGAGCTCTCTTTTGA 2245  
 Qy 2229 CAACCGTAACCTAAGATCCCTCAACTTGGAGAGCACCACCTGAAGAAAGATGATG 2288  
 Db 2246 CAATCAAACTTAAAGTACTCTCAATCTAGGAGACCTCCATGAAGATGATGATGATA 2305  
 Qy 2289 GATGCGGTGAGAGCTTAAACACCCAAATGTTGTGAGTCTTTGAGGCTGATG 2348  
 Db 2306 GTTAGCTGCGAAGGCTGAACATCAAGTGTCCGAGAGACTGAGGTTGATGATTC 2365  
 Qy 2349 CTGTGAGATTAACCATGATGTTATGATGATGATGATGATGATGATGATGATGATGATG 2408  
 Db 2366 CTGTGAGTAAACCATGATGTTATGATGATGATGATGATGATGATGATGATGATGATG 2425  
 Qy 2409 CCTGAATCTCTGAGCTGCGAGAAACAGGTTGACAGACAGGAGTAAATGCTCTGAG 2468  
 Db 2426 GCTAAAGTGTCTCAGCTGCGCAAAATGAGTGGAGTAAAGATGATGATGATGATGATG 2485  
 Qy 2469 TGAATGCTTTGAGAGTCTCCAGTGTGCGCTGCAAGCTGATCTGAGAGACTGTGAGAT 2528  
 Db 2486 GAATGCTTTGAGTACTCAATGTGTCTAAGTGAAGTGAATGATGATGATGATGATGATG 2545  
 Qy 2529 CACAGCCAGGCTTCCAGAGTCTGCTCAGCTCAGCTCTGTCAGCAACCGAGCTTGACCA 2588  
 Db 2546 CACACCTCCAGCTGCACTTGTGTCTGAGCTCTTTCAGCAACCAAGAACTTGACCA 2605  
 Qy 2589 CCTGTGCTTATCCAAACAGCTGAGGAGCAAGTGAATCTACGTGTGATGATGATGATGATG 2648  
 Db 2606 CCTGTGCTTATCCAAACAGCTGAGGAGCTGAGGAGTGAACAGCTGTGTGATGATGATG 2665  
 Qy 2649 GAGGCTTCCCACTGTGATCTGCAAGGCTGATGATGATGATGATGATGATGATGATGATG 2708  
 Db 2666 GAGGAATCAAGATGCTCTCAGCGCTGATGATGATGATGATGATGATGATGATGATGATG 2725  
 Qy 2709 TGGCTGTGTTTCTTGAATCTGCGCTTATGAGTAACTGATGATGATGATGATGATGATGATG 2768  
 Db 2726 TGTATGAGCTTCCGCAATGAGCTTGCAACCAACCAAGCTGACCCACCTGAGCT 2785

Qy 2769 TAGCATGAACCTGTGAGACATGAGCGTGAAGCTTCTGTGCGAGGTCAAGAGAAC 2828  
 Db 2786 GACCATGAACCCGTAGGAGATGAGTGAATGAAGCTATGTGTGAACCTTTAAAGAAC 2845  
 Qy 2829 ATCTTGTATCTCCAGACCTGAGATGTTGAAGTGTATCTCAACCCCGCTGCTGTA 2888  
 Db 2846 TACTTGTATCTTCAAAACATGGAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2905  
 Qy 2889 GAGTGTGCTGTGTGATCTCGAGAGACAGACCTGAAGGCTGTGATCTCAAGCA 2948  
 Db 2906 GGACCTGCTGTATGATCAACACACAGCACTTAAAGTTGATGATGATGATGATGATGATG 2965  
 Qy 2949 TGCCCTGGGTGACCGGTGGTGTGCTGCTGCGAGGAGCTGAAGCAAAAGAACAGTGT 3008  
 Db 2966 CGCCCTGGGTGACAAAGAGTCAATACCTGTGTGAGAGGAGCTGAAGCAAGTGAAGTGAAG 3025  
 Qy 3009 TCTGACGAGACTCGGTTGAAGGCTGAGCTGAATCTGATTTGCTGTGAGGACTCTC 3068  
 Db 3026 CTTGAGGAGACTTGGTGGGGGCTGAAGTTGATCTTCAATGCTGTGAGGACTGCTC 3085  
 Qy 3069 CTGGCCCTTCTCGCAACCGGACCTGACCAAGTCTAACTGTGTGAGAAATTAACCTTCAG 3128  
 Db 3086 ATGAGCAATCTTGGCAACCTCACTGAACAGCTAAACCTGTGAGAAATGATGATGATG 3145  
 Qy 3129 TCCAAAGAAATGATGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3188  
 Db 3146 TACATGGGAGATGTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3205  
 Qy 3189 AATTGGCTGTGAAATGACATGACCTGTGCAAAATGAAGAGTCTGAGAGAGTGA 3248  
 Db 3206 AATTGGCTGTGAAAGAGAGGATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3265  
 Qy 3249 GCTACTGAAGCCCGAGTGTGAATGAGCGATGATGATGATGATGATGATGATGATGATGATG 3308  
 Db 3266 GTTGTGAAGCCCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3325  
 Qy 3309 ACAC 3312  
 Db 3326 AAAC 3329

RESULT 14  
 AAD49001  
 ID AAD49001 standard; cdna; 3447 BP.  
 AC AAD49001;  
 XX  
 XX 07-MAR-2003 (first entry)  
 DT  
 XX  
 DE Mouse MATER cdna.  
 XX  
 XX Mouse; MATER protein; infertility; fertility; contraceptive agent;  
 KM gene therapy; gene; ss.  
 OS  
 XX  
 XX Mus sp.  
 FT  
 FT  
 FT Key  
 FT CDS  
 9..3344  
 Location/Qualifiers  
 /tag= a  
 /product= "Mouse MATER protein"  
 /transl\_except= (pos:648..650, aa:Thr)  
 /transl\_except= (pos:714..716, aa:Thr)  
 /transl\_except= (pos:768..770, aa:Thr)  
 /transl\_except= (pos:882..884, aa:Thr)  
 /transl\_except= (pos:1359..1361, aa:Thr)  
 /transl\_except= (pos:1554..1556, aa:Thr)  
 /transl\_except= (pos:1740..1742, aa:Thr)  
 /transl\_except= (pos:1869..1871, aa:Thr)  
 /transl\_except= (pos:2046..2048, aa:Thr)  
 /transl\_except= (pos:2049..2051, aa:Thr)  
 /transl\_except= (pos:2229..2231, aa:Thr)  
 /transl\_except= (pos:3216..3218, aa:Thr)



QY	2049	CATGCTGGGCAACCAACCCCAACACCTGGCGGACCTGGACCTGGGGGAGACGATCCGACAGA	2108
Db	2066	TGTGTGTTGG3CAGCCTCCGGAACTTGAAAGACCTGACCTTGG3GCAACAGCATCCGAGTCA	2125
QY	2109	GGGGGCCATGAAAGAACCTGTGTGCCAAGCTGAGGACATCCACCTGGCAAGATACAGACCCCT	2168
Db	2126	ACGGGCCATGAAAGATCTGTGCTCGAGCTCGGAMTACATGCTCTGGAGAAATACAGAAAGCT	2185
QY	2169	GATGTTTAGAAATGACAGATTAACCCCTGGTGTGACAGCACTCTGGAGAAATGTCATAGCC	2228
Db	2186	GACGTTTAAAGTGCAGAGGTAGTGTGCTGGCCCTGAACATCTCTGAAAGCTCCCTTTTTAG	2245
QY	2229	CAACCCGTACCTTAAGATTCCTCCACCTTGGAGGACCCACCTGAAGGAAGAGATGTAAAG	2288
Db	2246	CAATCAAAACTTAAAGTACCTCAATCTAAGGAAACCTCCATGAAAGGATGATGACATGAA	2305
QY	2289	GATGGCGGTGAAGGCTTAAACACCCAAATGTTTGTGGAAGTCTTGAAGGCTGAGATYG	2348
Db	2306	GTTAGCCCTGGAGAGGCTGAAACATCCAAATGCTCCGTGAGAGCTCTGAGGTTGAGATTC	2365
QY	2349	CTGTGAGATYGAACCAATGCTGTACCTGAAAGATCTCCCAAAATCCCTTACGACCTCCCCAG	2408
Db	2366	CTGTGAGTTAACATCATCTTGGTTATGAGATGATCTCCACGCTCTTATTTCAACACACAG	2425
QY	2409	CCTGAAATCTCTGAGCCTGGGCAAGAAACAAAGTGCACAGCAAGGAGATATGCTCTCAG	2468
Db	2426	GCTAAAGTGTCAAGCCTGGCCCAAAATGAGTGGAGTAAAAAGCATGATATCCCTTGG	2485
QY	2469	TGATGCTCTTGAAGAGTCTCCCAAGTGGCCCTGACAGAGGTGATATCTGAGAGCTGTGGAT	2528
Db	2486	GAAATGCTTGAAGTACCTAATGTGTCTACTGCAAAAGTTGATATCTGACAACTGTGGCT	2545
QY	2529	CACAGCCACGGGTTGCGAAGTCTTGACCTCAGCCCTGCTGACAAACGGAGCTTGAACA	2588
Db	2546	CACACCTGCAGCTGCACACTTCTGTGTCTGAGCCCTTTTCAGCAACGAACTTGAACA	2605
QY	2589	CCTGTGCTTATCCAAACAACAGCTTGGGGAAGAAAGTAAATCTACTGTGTCCATCCAT	2648
Db	2606	CCTGTGCTGTCCAAACAACAGCTTGGGGAAGTGAAGAGTGCACAAGCTGTCTCAGTTTCT	2665
QY	2649	GAGGCTTCCCACTGTAGTCTGACAGAGGCTGATGTGAATCAGTGGCCACTTGAACAAGCC	2708
Db	2666	GAGGAATCCAGAAATGTGCTCTTCCAGCGGCTGATCTGAATACCTGCAACATTTAGATGA	2725
QY	2709	TGGCTGTGTTTTCTTGACCTTGCGCTTATGGGTACTCATGCTGACGACCTGAGCT	2768
Db	2726	TGCTTATGGCTTCCGGCAATGAGACTTGCAAACAACAAGCTGACCCACCTGAGCCT	2785
QY	2769	TAGCATGAACCTGTGGAAGACAAATGGCGTGAAGCTTGTGCGAGGTGCATGAGAGAAC	2828
Db	2786	GACCATGAACCCCGTAGGGAGATGGTGCAAATGAAGCTACTGTGTGAAGCTTTTAAAGAAC	2845
QY	2829	AATCTGTATCTCCAGGACCTGGAATTGGTAAAGTATCTCAACCGCGCGTCTGTGA	2888
Db	2846	TACTTTGTTACTTCAAGAACCTGGAACCTAGTGAATGAGCACTGCAACAGAACTGTGTGGA	2905
QY	2889	GAGTCTGTCTGTGTGATCTCGAGAGGACAGACCTGAAGAGCTGTGATCTCAACGACAA	2948
Db	2906	GGACCTGGCTGTATGATCACAAACAACAACGACTTAAAGTTTGGATCTTGGTATCAAA	2965
QY	2949	TGCGCTGGGTGACGCTGGGGTGTCTGCGCTGTGCGAGGACTGAAGCAAAAGAACGTGT	3008
Db	2966	CGCCCTGGGTGACAAAGGAGTCAATTAACCTGTGTGAGGAGCTGAAGCAAAATGACAGCTC	3025
QY	3009	TCTGACGAGACTCGGGTGTGAAGGACATGACATACTTGTGATTTGTGTGTGAGGACATCTC	3068
Db	3026	CCTGAGGAGACTTGGGTGTGGG3GCAATGTAAGTTGACTTCAATGTCTGTGAGCAATGTCTC	3085
QY	3069	CTTGGCCCTTCTCTGCAACCGGCAATCTACCAAGTCTTAAACCTGGTGCAAGATTAATCTCAG	3128
Db	3086	ATTGGCCATCTCTTGCACACCTCAACCTGAACAGCTTAAACCTGGTGAAGATGACTTCAG	3145
QY	3129	TCCCAAGGAATGATGAAGCTGTGTTGGGCTTTGCTGTCCACGTCCTTAACCTTACAGAT	3188

Db	3146	TACATCGGGAGTGTGAAGCTGTGCTCGTTCCAATGCCCTGTCTCTAACCTGGGGAT	3205
Oy	3189	AATTCGGGCTGTGGAAATGAGCAGTACCCCTGTGGCAATTAAGAAAGCTGCTGGAGAGTGC	3248
Db	3206	AATTGGCGCTGTGGAAAGCAGGAGTACTATGCCCCGAGTGAGAAACAGCTGGAGGAATTGA	3265
Oy	3249	GCTACTCAAGCCCCAGACTCGTAATATGACCGTAGTGGCAATCTTTTATGAAATGACCG	3308
Db	3266	GTTTGTCAAGCCACCGTGTGATGTGATGGTATGTGATGATGATGAAATGACCG	3325
Oy	3309	ACAC 3312	
Db	3326	AAAC 3329	
RESULT 15			
ID	ABK48609		
XX	ABK48609	standard; CDNA; 1157 BP.	
AC	ABK48609;		
XX			
DT	13-AUG-2002	(first entry)	
XX			
DE	Human MATER CDNA fragment #1.		
XX			
KW	Human; gene; ss; contraceptive; antiinfertility; MATER;		
KW	Maternal antigen that embryos require; MATER null phenotype; oocyte;		
KW	early embryonic survival; premature ovarian failure; POF;		
KW	autoimmune infertility; chromosome 19; gene therapy; fertility.		
XX			
OS	Homo sapiens.		
XX			
XX			
XX			
Key	Location/Qualifiers		
FT	1. 1155		
FT	/tag= a		
FT	/product= "Human MATER"		
FT	/partial		
FT	/note= "No start or stop codon shown"		
XX			
PN	WO200232955-A1.		
XX			
PD	25-APR-2002.		
XX			
PF	04-APR-2001; 2001WO-US010981.		
XX			
PR	18-OCT-2000; 2000US-0241510P.		
XX			
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICES.		
XX			
PI	Nelson LM, Tong Z;		
XX			
DR	WPI; 2002-454595/48.		
DR	P-PSDB; AAU79523.		
XX			
PT	New isolated human Maternal Antigen That Embryos Require protein and		
PT	polynucleotide, useful in diagnosing or creating fertility or reduced		
PT	fertility, or as a contraceptive.		
XX			
PS	Claim 11; Page 68-69; 93pp; English.		
XX			
XX			
CC	The invention discloses an isolated human MATER (Maternal Antigen That		
CC	Embryos Require) protein, which can complement a MATER null phenotype in		
CC	which zygotes arising from the oocyte do not progress beyond the two-cell		
CC	stage. MATER is required for early embryonic survival and abnormal levels		
CC	of the protein can lead to premature ovarian failure (POF) and can be		
CC	caused by under or over expression of MATER or an autoimmune response to		
CC	MATER. MATER is a single-copy Maternal effect gene found on chromosome		
CC	19. The MATER protein and polynucleotide, by gene therapy, are useful in		
CC	diagnosing or treating fertility and reduced fertility. In particular,		
CC	the MATER protein is useful as a contraceptive agent, or for influencing		
CC	(either inhibiting or enhancing) fertility and can be used to detect a		
CC	predisposition to infertility or reduced fertility, or for presymptomatic		



screening of an individual for infertility/reduced fertility. The protein and polynucleotide are also useful for detecting an excess or deficiency, or genetic mutation, of the MATER protein in a mammalian subject (e.g. a human or a mouse) or for screening for a compound useful in influencing CC/MATER-mediated fertility. The sequence presented is the human MATER CDNA fragment #1, which was isolated from a human CDNA library

XX Sequence 1157 BP; 240 A; 324 C; 320 G; 273 T; 0 U; 0 Other;

Query Match 28.6%; Score 1152.8; DB 6; Length 1157;  
Best Local Similarity 99.8%; Pred. No. 9.3e-308;  
Matches 1154; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 754 CAGGCTCCGGTACCGAGATCATGTCGCCGACAGAAAGCTTTGTCATCTTACAGT 813
DB 1 CAAAGTCGGGTACGAGATCATGTCGCCGACAGAAAGCTTTGTCATCTTACAGT 60
QY 814 TTCGATGACCTGGGCTCTCTCTCAACATGACACAAAGCTTGCAGAGCTGGGCTGAG 873
DB 61 TTCGATGACCTGGGCTCTCTCTCAACATGACACAAAGCTTGCAGAGCTGGGCTGAG 120
QY 874 AAGCAGCCTCCGTTCAACCTCATACGAGTCGTGAGAGAGTCTCTGCTTCCCTGAGTCC 933
DB 121 AAGCAGCCTCCGTTCAACCTCATACGAGTCGTGAGAGAGTCTCTGCTTCCCTGAGTCC 180
QY 934 TTCCTGATGTCACCGTCAAGACGTCGGGCAACAGAAAGCTCAAGTCAGAGTCTGTCT 993
DB 181 TTCCTGATGTCACCGTCAAGACGTCGGGCAACAGAAAGCTCAAGTCAGAGTCTGTCT 240
QY 994 CCCCGTTACCTGTAGTGAAGAAATCTCCGGGGAACAAAGATCACTTGCTCTCTGAG 1053
DB 241 CCCCGTTACCTGTAGTGAAGAAATCTCCGGGGAACAAAGATCACTTGCTCTCTGAG 300
QY 1054 CGCGGGAATGTGAGCATCAGAAACACAAAGGTTGCGTGCATCATGAAACACCGTAG 1113
DB 301 CGCGGGAATGTGAGCATCAGAAACACAAAGGTTGCGTGCATCATGAAACACCGTAG 360
QY 1114 CTGCTCGACCAAGTCAGAGTCCCGCCGTCGTGAGTCTCTATCTGCTGGGCTCTGAGCTG 1173
DB 361 CTGCTCGACCAAGTCAGAGTCCCGCCGTCGTGAGTCTCTATCTGCTGGGCTCTGAGCTG 420
QY 1174 CAGGACGTGTGGGGAGAGCGTCGCCCTTCAACCAAGCTCAACAGGCTGACAGGCC 1233
DB 421 CAGGACGTGTGGGGAGAGCGTCGCCCTTCAACCAAGCTCAACAGGCTGACAGGCC 480
QY 1234 GCTTTTGTGTTTCATCAGCTCAACCCCTGAGGCGTGTCCGGCGCTGTCTCATCTGAG 1293
DB 481 GCTTTTGTGTTTCATCAGCTCAACCCCTGAGGCGTGTCCGGCGCTGTCTCATCTGAG 540
QY 1294 GAAAGATTTGTCTGAAAGGCTTTCGCCGTATGCTGTGAGGAGTGTGAATAGGAAG 1353
DB 541 GAAAGATTTGTCTGAAAGGCTTTCGCCGTATGCTGTGAGGAGTGTGAATAGGAAG 600
QY 1354 TCAGTGTTCATGATGAGGAGCTCATGTTCAAGGACTCGGGAGTCTGAGCTCGGTGCT 1413
DB 601 TCAGTGTTCATGATGAGGAGCTCATGTTCAAGGACTCGGGAGTCTGAGCTCGGTGCT 660
QY 1414 CTGTTTCAATGAAATCTTCTCCAGACAGCCACTGTGAGAGTACTACACTTCTTC 1473
DB 661 CTGTTTCAATGAAATCTTCTCCAGACAGCCACTGTGAGAGTACTACACTTCTTC 720
QY 1474 CACTCAGTCTTCAGAGACTTCTGTGCGCGCTTGTACTACGTGTGAGGGCTGGAATC 1533
DB 721 CACTCAGTCTTCAGAGACTTCTGTGCGCGCTTGTACTACGTGTGAGGGCTGGAATC 780
QY 1534 GAGCAGCTCTCTGCGCTCTGATGTTGAGAGACAAAGAGTCCATGAGCTTAAACAG 1593
DB 781 GAGCAGCTCTCTGCGCTCTGATGTTGAGAGACAAAGAGTCCATGAGCTTAAACAG 840
QY 1594 GCAGGCTTCATATCACTCGCTTGTGATGAGAGCGTTCTTGTGCTCGTGAGCGAA 1653
DB 841 GCAGGCTTCATATCACTCGCTTGTGATGAGAGCGTTCTTGTGCTCGTGAGCGAA 900

```

```

QY 1654 GACGTAAAGAGGCCACTGAGAGTCTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAG 1713
DB 901 GACGTAAAGAGGCCACTGAGAGTCTGCTGGGCTGTCCCGTTCCCTGGGGGTGAAGCAG 960
QY 1714 AAGCTTTCAGACTGGGTCTCTGTGTTGGGTGAGCAGCCTTAATGCCACCCAGAGAGC 1773
DB 961 AAGCTTTCAGACTGGGTCTCTGTGTTGGGTGAGCAGCCTTAATGCCACCCAGAGAGC 1020
QY 1774 ACCCTGAGCGCTTCCACTGTCTTTTGAGACTCAAGACAAAGATTTGTCCTTGCA 1833
DB 1021 ACCCTGAGCGCTTCCACTGTCTTTTGAGACTCAAGACAAAGATTTGTCCTTGCA 1080
QY 1834 TTAACAGCTTCCAAAGAGTGTGGCTCCGATTAAACAGAACTGGACTTGATACATCT 1893
DB 1081 TTAACAGCTTCCAAAGAGTGTGGCTCCGATTAAACAGAACTGGACTTGATACATCT 1140
QY 1894 TCCTTCTGCTTCAGC 1909
DB 1141 TCCTTCTGCTTCAGC 1156

```

Search completed: July 18, 2005, 16:47:42  
Job time : 2052 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2005, 16:15:50 ; Search time 632 Seconds  
(without alignments) 10446.798 Million cell updates/sec

Title: US-10-066-521-5

Perfect score: 4035  
Sequence: 1 atggaagagagagaatcgct.....tcgaggggctggtgtctta 4035

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405566

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/lna/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/lna/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/lna/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/lna/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/lna/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/lna/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	221.4	5.5	4422	US-09-388-221B-1	Sequence 1, Appli
2	221.4	5.5	4556	US-09-388-221B-9	Sequence 9, Appli
3	220.6	5.5	4200	US-09-388-221B-3	Sequence 3, Appli
4	220.6	5.5	4332	US-09-388-221B-5	Sequence 5, Appli
5	220.6	5.5	4466	US-09-388-221B-11	Sequence 11, Appli
6	183.8	4.6	1371	US-08-910-731-1	Sequence 1, Appli
7	183.8	4.6	1371	US-08-910-731-5	Sequence 5, Appli
8	182.4	4.5	1386	US-08-910-731-7	Sequence 7, Appli
9	180.8	4.5	1372	US-08-910-731-3	Sequence 3, Appli
10	175.8	4.4	1374	US-08-910-731-3	Sequence 3, Appli
11	175.8	4.4	1374	US-08-910-731-3	Sequence 3, Appli
12	61	1.5	7218	US-08-232-463-14	Sequence 14, Appli
13	61	1.5	376	US-09-023-655-395	Sequence 395, App
14	51.4	1.3	298336	US-09-949-016-15600	Sequence 14600, A
15	50.8	1.3	64309	US-09-949-016-14581	Sequence 14581, A
16	47.4	1.2	4485	US-09-949-016-3209	Sequence 3209, Ap
17	47.4	1.2	4485	US-10-014-269-1	Sequence 1, Appli
18	47.4	1.2	4485	US-10-014-269-1	Sequence 33, Appli
19	46.6	1.2	1276	US-09-177-325-2	Sequence 2, Appli
20	46.6	1.2	1276	US-09-411-812A-2	Sequence 2, Appli
21	46.6	1.2	1276	US-09-590-113-2	Sequence 358, App
22	46.4	1.1	2780	US-09-620-312D-358	Sequence 14, Appli
23	45	1.1	929	US-09-671-317-14	Sequence 340, App
24	44.8	1.1	1620	US-09-134-000C-340	Sequence 439, App
25	43.4	1.1	1001	US-09-671-317-439	Sequence 37, Appli
26	43.2	1.1	193303	US-09-497-855A-37	Sequence 44, Appli
27	43.2	1.1	193303	US-09-497-855A-44	

28	42.8	1.1	1449	US-09-248-796A-4693	Sequence 4693, Ap
29	42.8	1.1	767677	US-09-949-016-12147	Sequence 12147, A
30	42.8	1.1	767677	US-09-949-016-11361	Sequence 11361, A
31	42.6	1.1	693	US-09-252-991A-11050	Sequence 11050, A
32	42.6	1.1	1005	US-09-252-991A-10976	Sequence 10976, A
33	42.6	1.1	1185	US-09-252-991A-11203	Sequence 11203, A
34	42.4	1.1	150394	US-09-949-016-13042	Sequence 13042, A
35	42.2	1.0	38566	US-09-949-016-15271	Sequence 15271, A
36	42.2	1.0	38566	US-09-949-016-15272	Sequence 15272, A
37	42	1.0	3117	US-09-614-221A-275	Sequence 275, App
38	41.8	1.0	7044	US-09-614-221A-1113	Sequence 14113, A
39	41.6	1.0	2223	US-08-257-073-4	Sequence 4, Appli
40	41.4	1.0	505	US-09-621-976-15639	Sequence 15639, A
41	41.2	1.0	810	US-08-742-185-103	Sequence 103, App
42	41.2	1.0	43795	US-08-742-185-101	Sequence 101, App
43	41	1.0	505	US-09-621-976-15639	Sequence 15639, A
44	41	1.0	10391	US-09-902-540-958	Sequence 958, App
45	40.4	1.0	240157	US-09-949-016-16264	Sequence 16264, A

## ALIGNMENTS

RESULT 1					
US-09-388-221B-1					
Sequence 1, Application US/09388221B					
Patent No. 6818750					
GENERAL INFORMATION:					
APPLICANT: Reed, John C.					
TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation					
FILE REFERENCE: P-IJ 3650					
CURRENT APPLICATION NUMBER: US/09/388,221B					
CURRENT FILING DATE: 1999-09-01					
NUMBER OF SEQ ID NOS: 30					
SOFTWARE: PatentIn Ver. 2.0					
SEQ ID NO 1					
LENGTH: 4422					
TYPE: DNA					
ORGANISM: Homo sapiens					
FEATURE:					
NAME/KEY: CDS					
LOCATION: (1)..(4422)					
US-09-388-221B-1					
Query Match					
Best Local Similarity 47.9%; Pred. No. 9.1e-55;					
Matches 953; Conservative 0; Mismatches 966; Indels 72; Gaps 8;					
QY	532	TTTGAATTCAGACCGGCTGCGGCTTCCGCTGCGACCGTGTTCGACGGAATCAGGA	591		
DB	952	TTTGCCCAAGGCTTGATATCCAGAACCTGCAATGATCACTGACAGGGGCTGCTGA	1011		
QY	592	ATTGGGAATGAGGCTCTGACGAGAGATCGTGTGCGGCGGAGTGAAGTCTTAC	651		
DB	1012	ATTGGGAATGAGACTGCGGAGAGTGAAGAGCTTGGGAGAGGCCAGCTGAT	1071		
QY	652	CAGGAATGCTCTCTACGCTCTTCTCTCCCGTTAGAGATGACAGCGGAAGAG	711		
DB	1072	GGGAGCGCTTCCAGCATGTCTTCACTTCACTGACAGAGCTGCGCCAGTCCAGG	1131		
QY	712	AGCAGTGCACAGATTATCTCCAGGAGTGGCCAGCTCCAGGCTCCGATACGAG	771		
DB	1132	GTGATCTGCTGAGCTCATGGAAGATGAGACAGCACTCCGCTCCATTAGACAG	1191		
QY	772	ATCATGTCACACAGAAAGGCTGTTGTCATCATATTCAGGTTTCATGACCTGGCTCT	831		
DB	1192	ATTCGTGTAGGACAGAGCGGCTCTTATCTTCATGATGATGATGACAGAGATG	1251		
QY	832	GTCTCAACAAATGACA---CAAAGCTGTCAAAAGCTGGCTGAGAGACGCTTCGTT	888		
DB	1252	GTCTTGCAGAGACGAGATTGAGCTGTGATGACAGACCAACGACGCGGAGAT	1311		
QY	889	ACCCTATACGACGATGCTGCTAGAGAGAGTCTGCTCCCTGAGTCTTCTATGCTACC	948		

```
Db      1312 GCACCTGCTGGGAGATTGCTGGGAAACATACTACCTCCGAGGATCCTTCTGTATCAG 1371
Qy      949  GTCAAGACGTGGGACACAGAAAGTCAAGTCAGAGTGTGTCTCCCGTTACTGTTA 1008
Db      1372 GCTCGGACCAAGCTCTGCAAGAACCTCATCTCTTGTGGAGCAGGCACTGTGGGTAAAG 1431
Qy      1009 GTTAGAGATCTCGGGGAAACAAGATCCATTGCTCTGAGCGGGATGTGGTAG 1068
Db      1432 GTCCGGGGTCTCTGAAGTCCAGCAGAAAGAAATTTTCTACAGATTTTCAAGATAA 1491
Qy      1069 CATCAAGACACAAGGTTGGTCGATCAATGAACAACGTGAGCTCTGACCAATGC 1128
Db      1492 AGGCAAGCAATTAAGACCTTTAGTTGATCAATCAAAAGAGCTCTGGGCCCTGTGT 1551
Qy      1129 CAGGTGCCCCCGTGGGCTCTCTCATCTGCGTGGCCCTGACGTGACAGAGCTGTGGG 1188
Db      1552 CTGTGCTCGGTGGTGTCTGTGGCTGGCTGCACTTGCCTGATGCAAGATGAAGCGGAAG 1611
Qy      1189 GAGAGCGTGGCCCCCTTCAACCAACGCTCAAGGCTGCAAGCGGCTTTTGTGTTCAT 1248
Db      1612 GAAAACTCACTGACTTCCAAAGACACACACACCTCTGTCTATCATTTACTTGGCCAG 1671
Qy      1249 CAGCTCACCCCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTGAGAAAGATTGTCTG 1308
Db      1672 GCTCTCAAGCTCAAGCATTTGG-----GACCCAGCTC 1704
Qy      1309 AAGCGCTTTCGCGGTATGCTGTGGAGGAGTGTGAATAGAGAGTCACTGTTGATGCT 1368
Db      1705 AAGACCTCTGTCTCTGTGCTGTGAGGCACTGTGCAAAAACACCTTTTCACTGCA 1764
Qy      1369 GACGACCTCATGTTCAAGSACTGGGGAGTCTGAGCTCCGTGCTGTGTTCACTGAC 1428
Db      1765 GATGACCTCAAGAGACATGGGTTAGATGGGCTCATCTTCCACTTTTGAAATGGGT 1824
Qy      1429 ATCTCTTCCAGACAGCCACTGTGAGAGATCTAACACCTTCTTCCACTGATCTCCAG 1488
Db      1825 A---TTCTTCAAGACACCCCAATCCCTCGAGCTACAGCTTCATTCACCTGTGTTTCA 1881
Qy      1489 GACTCTGTGCGCGCTTGTATCACTGTTTGAAGGCGCTGGAATATGAGCCAGCTCTGCG 1548
Db      1882 GAGTCTTTTGAAGCATGTCTCTATGTCTTGTGAGATGAAGAGGGAGAGTAAACATTTCT 1941
Qy      1549 CCTCTGTAAGTTGAAGACAAGAGTCAATGAGCTTAAACAGGAGGCTTCCATATC 1608
Db      1942 AATTGCATCATATGATTTGAAAAGCGCTAGAAGCATATGAAATCAGGCTGTT--- 1998
Qy      1609 CACTCGCTTTGAGTAAACGTTTCTTGTGGCTCTGAGAGGAGACGTAAAGAGCCCA 1668
Db      1999 ---GGGGCATCAACACAGCTTTCTATTTGGGCTGTTAATGATGAGGGGAGAGAGAG 2055
Qy      1669 CTGAGAGTCTGCTGGGCTGTCCGTTCCCTGGGGGTGAAGACAGAGCTTCTGCATCG 1728
Db      2056 ATGGAGAACATCTTTCATCTGCGGCTGTCTCAAGGGAGAAACCTGATCAGTGTGCTCCG 2115
Qy      1729 GTCTCTGTGGGTGACAGCGCTTAATGCAACACCCAGAGACACCTGGAAGGCTTC 1788
Db      2116 TCCCTGACGCTGCTGTGAGC-----ACACTCTCTGAGTCCCTC 2157
Qy      1789 CACTGTCTTTTGAAGTCAAGACAAAGATTGTTGCTTGGCACTTAAACAGCTTCCAA 1848
Db      2158 CACTGCTTTGAGAGACTCGAGAAAGAAAGTTTCCGACACAAAGTATGGCCCATTTGAA 2217
Qy      1849 GAAGTGTGGCTTCGATTAAACAGAACTGACTTGAATGACATCTTCTTGTGCTCCAG 1908
Db      2218 GAAATGGGCACTGTGTGAAACAGACATGGAGCTTATGTGTGACATTTTCTGCACTTAA 2277
Qy      1909 CACTGTCCGTATTTGGGAAATTCGGGTGATGATCAAGATTTCCCAAGATAG 1968
Db      2278 TTCAAGCCCAAGTAAAGCTTGAAGTGAATGAGGAGG-----CAGCAAGATC 2330
Qy      1969 TCCGCTGAGGAGATGCTGTGTGCTCTATGATGAGCGGAGTAAAGACCTCATTTAGAGAG 2028
```

```
Db      2331 AACATGAGCCCCCAGCATGTGATGCTGTTCAGGTGGGTCC-----CAGTCAAGATGCC 2385
Qy      2029 CAGTGGAAAGATTTCTGTCTCCATGCTTGGACCCACCAACCTTGGCGAGCTGAGACTG 2088
Db      2386 TATTGGAGATTTCTTCTCTGCTCCAGAGTCAACGAAACCTGAAGAGCTGGACCTTA 2445
Qy      2089 GGCAGCAGATCTCGACAGAGCGGGCCATGAAGACCTGTGTGCAAGTGAAGGACATCCC 2148
Db      2446 AGTGAAGATCGCTGAGCCACTGTGCAAGTGAAGAGTCTTTGTAAAGACCTTGAAGCGCCCT 2505
Qy      2149 ACCTGCAAGATACAGACCTGTATGTT--TAGAAATCAAGATTAACCTGTGTGTCAG 2205
Db      2506 CGTGCCTCCCTGAGAGACCTGCGGTGAGCTGTGCTGTGCTCAACAGCTGAGAGTGAAG 2565
Qy      2206 CACTCTGAGAAATCGTCAATGCGCAACCGTAAACCTAATGCCCTCAACCTTGGAGGAC 2265
Db      2566 GACTTGTCTTTGGGCTGAGAGCCACACAGACCTTGAACGAGCTGAGCTGAAGCTTCAAT 2625
Qy      2266 CACTGAAGAGAGAGATGTAAAGATGCGGTGAACCTTAAACACCCAAATGTTTG 2325
Db      2626 GTGCTCAAGATGCTGAGAGCCAAACACTTTGCCAGAGACTGAGACAGCCGAGCTGCAAG 2685
Qy      2326 TTGAGTCTTTGAGCTGATGCTGTGTGATTAACCATGCTGTAACTTGAAGATCTCC 2385
Db      2686 CTACAGGAGACTGACAGTGTGCTGTGAGCTGTGCTCACTGCTGTGCTGCAAGACTGAGCC 2745
Qy      2386 CAATCTTACAGACTCCCGAGCGCTGAATCTGTGAGCCCTGAGAGGAGAAACAAGTGA 2445
Db      2746 TCTGTGCTTATGTGCAACCCCGAGCTGAAGAGCTGAACCTGACAGAAACACTGTGAT 2805
Qy      2446 GACAGGAGATTAATGCTCTCAAGATGCTTGAAGTCTTCCAGCTGCGCCCTGACAAG 2505
Db      2806 GACTTGGCGTGGCAGCTGCTGTGAGGGGCTCAGGAGCTGTGCTGTCAAACTATACGC 2865
Qy      2506 CTGATTAAGTGA 2516
Db      2866 CTGGGCGTGA 2876
```

```
RESULT 2
US-09-388-221B-9
; Sequence 9, Application US/09388221B
; Patent No. 6818750
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221B
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 4556
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4365)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Construct
US-09-388-221B-9
```

```
Query Match 5.5%; Score 221.4; DB 4; Length 4556;
Best Local Similarity 47.9%; Pred. No. 9.3e-55;
Matches 953; Conservative 0; Mismatches 966; Indels 72; Gaps 8;

Qy      532 TTGATTCAGACCGGCTGGGCTTCCGCTCGCAAGGTGTTTGCACGGAATTCAGGA 591
Db      952 TTTGGCCCAAGGCTGGATTAACCAAGAACTTCGCTATGCTACATGCAAGGGGCTGCTGGA 1011
Qy      592 ATTGGAAATCGGCTTACGACAGAAAGATCGTGTGTGCTGGGCGCAAGGTGACTATAC 651
```

Db 1012 ATTGGGAATCAACATGCGCAGAGCAGGTAAGAAAGCCTGGGGAGAGGCCAGCTGTAT 1071  
Qy 652 CAGGGAATGTTCTCTACGCTCTTCTCCCGTTAGAGATGACAGGGAAGAGAG 711  
Db 1072 GGGGACCCCTTCACATGTCCTTCTACCTGACGTCAGAGAGCTGCCCCAGTCAGAGG 1131  
Qy 712 AGCATGTCACAGATTCTCCAGGGAGTGGCCAGACTCCAGGCTCCGGTGAAGAG 771  
Db 1132 GTGAGTCTGCTGAGACTCATCGAAAAAGTGGAGACGCACTCCGGCTCCATTAGACAG 1191  
Qy 772 ATCATGTCCCAACAGAAAGGCTGTGTTGATCATTTAGACGTTTCGATGACCTGGGCTCT 831  
Db 1192 ATTCGTGTAGGCGCAGAGGGGCTGCTCTCATCTCGATGTGTAGATGACAGAGATGG 1251  
Qy 832 GTTCCCAACATGACA---CAAGCTCTGCAAAAGCTGGGCTGAGAGAGAGCTCCGTTCC 888  
Db 1252 GTCTTGACAGAGCCAGATTGAGCTGTGTCTGCACTGAGACCAACAGCCGCGGAT 1311  
Qy 889 ACCCTCATAGCAGCTGTGCTGAGAAAGGCTGTCTCCCTGAGTCCCTGATGCTCAC 948  
Db 1312 GCACGTCTGGGAGATTGTGCTGGGAAAACTATCTCCGAGGCACTCTTCTGATCACG 1371  
Qy 949 GTCAGAGACGTGGGACAGAGAAAGCTCAAGTCAGAGTGTGTCTCCCGTTACCTGTTA 1008  
Db 1372 GCTCGGACCAAGCTCTGAGAACTCAATCTTCTTGAGACAGGCAAGTTGGGTAGAG 1431  
Qy 1009 GTTAGAGAAATCTCCGGGAAACAAAGATCCACTGTCTCTTGAGCGGGATTTGGTGA 1068  
Db 1432 GTCTGGGGTCTCTGAGTCCAGAGAAAGAAATTTCTACAGATTTTCAACATGAA 1491  
Qy 1069 CATCAGAGACAAAGGGTGTGCGGATGAAACCGTGAAGCTGTGACAGATGC 1128  
Db 1492 AGGCAAGCAATTAAGACCTTTAGTTGGCTCAATCAACAAAGACCTCTGGGCTGTGT 1551  
Qy 1129 CAGGTCCCGCGGTGCTCTCTCATCTGCGTGCCTGACAGCTCAGAGCTGTGGG 1188  
Db 1552 CTGTGGCTCTGGGTCTCTGGCTGGGCTGCACTTGCCTGATGACAGATGAAGGGAAG 1611  
Qy 1189 GAGAGCGTGGCCCTTCAACCAAGGCTCAGAGGCTGACGCGCTTTGTGTGAT 1248  
Db 1612 GAAAAAATCAACATGACTTCCAAAGACCAACACCTCTGTCTAATTAACCTTGCAG 1671  
Qy 1249 CAGCTCACCCCTCAGAGGCGGTGTCGAGGCTGTCTCAATCTGAGAGAAAGTTGCTCG 1308  
Db 1672 GCTCTCAAGCTCAGCCATTGG-----GACCCAGCTC 1704  
Qy 1309 AAGCGCTTCTGCGGTATGCTGTGAGGAGAGTGTGAATAGGAATGATGTTGATGT 1368  
Db 1705 AGAGACCTCTGTCTCTGCTGTGAGGCACTGGCAAAAAGAACCTTTTTCAGTCA 1764  
Qy 1369 GACGACTCATGTTCAAGAGACTCGGGAGTCTGAGCTCGTCTGTTTCAATGAAC 1428  
Db 1765 GATGACCTCAGAGACATGAGTGTAGATGGGCACTATCTCCACTTCTGTAAGATGGAT 1824  
Qy 1429 ATCTCTTCCCAAGACGACGACGTGAGAGATCAACCTTCTTCACTCACTGCTCAG 1488  
Db 1825 A---TTCTTCAAGACACCCCATCTCTGTGAGCTACACTTCACTCTGTTTCAA 1881  
Qy 1489 GACTTCTGTGCGCGCTTGTACTACGTGTTAGAGGCGCTGAAATCGACGACGCTCTGC 1548  
Db 1882 GAGTTCCTTGTGAGCAATGTCTTATGTCTTGAGAGATGAGAAAGGGAAGAGTAACTTCT 1941  
Qy 1549 CCTCTGATGTTGAGAGACAAAGAGTTCATGAGCTTAAACAGGAGCTTCCATATC 1608  
Db 1942 AATTGCATCATGATTTGAAAAAGCGCTAGAGCATGTGAATCATGAGCGCTGTT--- 1998  
Qy 1609 CACTGCTTTGGATGAAGCGTTTCTGTGTTGGCCCTGAGAGAGAGAGATGAAGAGGCA 1666  
Db 1999 ---GGGGATCAACACAGTTCCTATTTGGGCTGTTAAGTGAAGGGGAGAGAG 2055  
Qy 1669 CTGAGAGTCTGTGGGCTGTCTCCGCTGGGGGTGAAGCAGAACTTCTGCACTGG 1728  
Db 2056 ATGGAAGACATCTTCACTGCGGCTGTCTAGGGGAGAACTGATGACATGGGTCCG 2115

Qy 1729 GTCTCTGTGGGCTCAGAGCTTAATGCAACCAACCCAGAGACACCTGAGCGCTTC 1788  
Db 2116 TCCCTGACAGCTGTGCTCAGC-----ACACTCTCTGAGATCCCTC 2157  
Qy 1789 CACTGTCTTTTTCGAGACTCAAGACAAAGATTTGTTCCGCTTGGCAATTAACAGCTTCAA 1848  
Db 2158 CACTGCTTTGAGAGACTTCGGAACAAAGCTTCTGACACAGATGATGGCCATTTCGAA 2217  
Qy 1849 GAAGTGGCTTCCGATTAAACAGAACTGGAATTGATGACATCTTCTTCTGCTCAG 1908  
Db 2218 GAATGGGCAATGTGTGTAAGAAACAGACATGAGAGCTTATGTGCACTTTCGCAATTA 2277  
Qy 1909 CACTGTCCGTAATTTCCGAAAAAATTCGGGTGATGTCAAAGGATTTCCCAAGATGAG 1968  
Db 2278 TTCAAGCCGCAAGTAAGAACTTCAAGCTGATGAGGCAAG-----CAGCAAGATC 2330  
Qy 1969 TCCGCTGAGGCAATGCTGTGCTCCCTTAATGATGCGGATTAAGACCTCAATTGAGAG 2028  
Db 2331 AACTAGAGCCCAACATGATGATGCTGTGATGAGTGGTCC-----CAGTCAAGATGCC 2385  
Qy 2029 CAGTGGGAAGATTTCTGCTCAATGCTTGGACCAACCAACCTGCGGACGCTGACCTG 2088  
Db 2386 TATTGGCAGATTTCTTCTCCGCTCTCAAGGTCAACAGAAACCTGAAGAGAGCTGGA 2445  
Qy 2089 GGCAGACATCTCTGACAGAGAGGCGCATGGAAGACCTGTGTGCCAAGCTGAGCATGCC 2148  
Db 2446 AGTGAACCTGCTGAGCACTCTGCAAGTGAAGTCTTTGTAAAGACCTTGAGAGGCTT 2505  
Qy 2149 AACTGCAAGATTAAGACCTCTGATGT---TAGAAATGACAGATTAACCTGTGTGAG 2205  
Db 2506 CGCTGCTCTGAGACCTTCCGTTGGCTGTGCTGTGTGCTCTCAAGCTGAGAGCTGCAAG 2565  
Qy 2206 CACTCTGAGAAATGCTGATGAGCAACCGTAACCTTAAGATCTTCACTTGGAGGCA 2265  
Db 2566 GACTGTGCTTTGGGCTGAGAGCCCAACAGACCTGACCGAGCTGAGACTTGAAT 2625  
Qy 2266 CACTGAGAGAGAGATGTAAGATGCGGTGTGAAGCTTAAACACCCAAATGTTTG 2325  
Db 2626 GTGTCAAGAGATGCTGAGGCAACCACTTTGCCAGAGATGAGACAGCCGAGCTGCAAG 2685  
Qy 2326 TTGAGTCTTTGAGGCTGATTTGCTGTGATTTGACCCATGCTGTACTGTAAGATCTC 2385  
Db 2686 CTACAGGACATGCAAGCTGTGTGAGCTGTGCTCTCACTGTGCTGTGCTGACAGACTG 2745  
Qy 2386 CAAATCTTACAGCTCCCGCAGGCTGAATCTGTGAGCTGTGAGAGAAACAGGTGCA 2445  
Db 2746 TCTGTGCTTATGTCAGAGCCCGCAGCTGAAGAGCTGACCTGACAGACAGAACTGAT 2805  
Qy 2446 GACCAAGGAGTAAATGCTCTGATGATGCTTGAAGATCTTCCAGTGGCGCTGACAG 2505  
Db 2806 GACGTTGCGTGGCACTGCTGTGAGGGGCTCAGGCACTCTGCTGCAACTATACGC 2865  
Qy 2506 CTGATACTGA 2516  
Db 2866 CTGGGCTGGA 2876

RESULT 3  
US-09-388-221B-3  
; Sequence 3, Application US/09388221B  
; Patent No. 6818750  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation  
; FILE REFERENCE: P-1J 3650  
; CURRENT APPLICATION NUMBER: US/09/388,221B  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 4200  
; TYPE: DNA

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(4197)  
US-09-388-221B-3

Query Match 5.5%; Score 220.6; DB 4; Length 4200;  
Best Local Similarity 48.0%; Pred. No. 1.5e-54;  
Matches 939; Conservative 0; Mismatches 944; Indels 72; Gaps 8;

QY 532 TTTGATTCAGACCGGTGGGGCTTCCGGCTTCGACCGGTGTTCTTCGACGGAAGTCAGGA 591  
DB TTTGGCCACGCGCTGATATCCCAAGAACTCGCATATGCAATACGACGGGGCTGCTGGGA 1011  
QY 592 ATTGGGAAATGGGCTTAGCCAGAAAGATCGTGTGCTGTGCTGGGGGCAAGGTGGACTGTAC 651  
DB ATTGGGAAATGGGCTTAGCCAGAAAGATCGTGTGCTGTGCTGGGGGCAAGGTGGACTGTAC 1071  
QY 652 CAGGAAATGTTCTCTCAAGTCTTCTCTCCCGTTAGAGAGATGACGCGAAGAGAGAG 711  
DB GGGGACCGCTTCCAGCATGTCTTCTACTTCAGTGCAGAGAGCTGGCCCATGCCAAGGTG 1131  
QY 712 AGCAGTGCACAGATTATCTCAAGGAGTGGCCAGACTCCAGGCTCCGTTACGAG 771  
DB GTGAGTCTGCTGAGCTCATCGAAAGATGGGACAGCACTCCGGCTCCCATTTAGACAG 1191  
QY 772 ATCATGTCGCGACGAGAAAGGCTGTGTTTCATCATGAGGTTTCGATGACCTGGGCTCT 831  
DB ATCTGTGTTAGGCGAGAGCGGCTGCTCTTCACTTCGATGATGATGATGATGATGATGATG 1251  
QY 832 GTCTCAACAATGACA---CAAAGCTTCGAAAGACTGGGCTGAGAGACAGCTCCGTTCC 888  
DB GTCTTGGAGAGAGCGAGTTCGAGCTGTCTGCACTGGAGCCAGCCAGCCAGCCGCGGAT 1311  
QY 889 ACCCTCATACGAGTCTGCTGAGGAAGTCTGCTCCCTGAGTCTTCTGATGTGACCC 948  
DB GCACGTGCTGGGCGAGTTGCTGGGGAATACTACTCCGAGGCACTCTTCTGATACAG 1371  
QY 949 GTCAAGAGCGTGGGACAGAGAGCTCAAGTCAAGGCTGCTGCCCGTTACTCTGTTA 1008  
DB GCTGGACACAGCTCTGCAAACTCATTCCTTCTTGGAGCAGGACAGCTGGGTTAGAG 1431  
QY 1009 GTTAGAGAAATCTCCGCGGAAACAAAGATCCACTTGTCTTGAAGCGCGGAGTTGTAG 1068  
DB GTCCGGGGTTCCTGAGTCCAGAGAGGAATATTTCTACAGATATTTTACAGATGAA 1491  
QY 1069 CATCAAGAGACAAAGGTTGCTGTCATGAAACACCTGAGCTGCTGACCAAGTGC 1128  
DB AGGCAAGCAATTAAGGCTTTAGGTTGTTCAATCAAAACAAAGAGCTCTGGGCCCTGTGT 1551  
QY 1129 CAGGTCGCCGCGTGGGCTCTCATCTGCGTGGCCGAGCTGACAGAGAGTGGTGGG 1188  
DB CTGTGGCTGGGGTCTGTGGCTGGCTGACCTGCTGATGACAGCATGAGGGAG 1611  
QY 1189 GAGAGCGTGGCCCTTCAACCAACGCTCAAGGCTGCAAGCGGCTTTTGTGTTTCA 1248  
DB GAAAACTCACTGACTTCAAGACACCAACCTCTGTCTACATTTACCTTGGCCAG 1671  
QY 1249 CAGCTCAACCTCGAGGCGTGGTCCGGCGCTGTCTCAATCTGAGAGAAAGTTGTCTG 1308  
DB GCTCTCCAAAGCTCAAGCCATTGG-----GACCCAGCTC 1704  
QY 1309 AAGGCTTCTGCGGATGCTGTGGAGGAGGTGGAATAGGAAGTCAGTGTGATGATG 1368  
DB AGAAGCTCTGCTCTGTGGCTGAGGGCATCTGGCAAAAAAAGACCTTTTCACTCA 1764  
QY 1369 GACGACCTCATGTTCAAGAGACTCGGGAGTCTGAGCTCCGTTCTGTTCATCATGAAAC 1428  
DB GATGACCTCAAGAGAGAGGTTGATGGGGCATCATCTCAACCTTTGAAGTGGGT 1824  
QY 1429 ATCTCTTCCCAAGAGCACTGTGAGAGATCACTACCTTCTTCACTGATCTCAG 1488  
DB A---TTCTTCAAGAGCAACCCATCTCTGAGCTACAGCTTCACTTCACTCTGTTCAG 1881

QY 1489 GACTTGTGCGCGCTTGACTACTAGTGTAGAGGGCTGGAAATCGAGCCAGCTCTCTGC 1548  
DB GAGTTCCTTTGACGAATGTCTTATGTCTTGGAGATATGAAAGGGAGAGTAACATCT 1941  
QY 1549 CCTCTGATGTTGGAAGACAAAGAGGTCCATGAGACTTAAACAGGAGCTTCCATATC 1608  
DB AATTGACATATGATTTGAAAAGACGCTAGAAAGCATATGAAATACATGAGCTGTGGG 2001  
QY 1609 CACTGCTTTGATGAAAGCTTTCTGTGTGGCTCGTGAAGCCGAAAGCTTAAGAGGCCA 1668  
DB GCATC-----AACCAAGCTTTCATATGAGGCTGTAAAGTGAAGGGAGAGAGAG 2055  
QY 1669 CTGAGAGTCTGCTGGGCTGTCCCGTTCCCTGGGGGTGGAAGCAGAGCTTCTGACCTGG 1728  
DB ATGAGAAACATCTTTCATCTGCGGCTGTCTCAGGAGAGAACTGATGCTGTGCTCCG 2115  
QY 1729 GTCTCTGTGGGTGACAGGCTTAATGCAACACCCAGAGAGACCTTGAAGCTTC 1788  
DB TCCTTGAGCTGTGCTGACGCT-----ACACTCTGAGAGTCCCTC 2157  
QY 1789 CACTGTCTTTTCCAGACTCAAGACAAAGATTTGTGCTTGGCATTTAAACAGCTTCAA 1848  
DB CACTGCTTTGACGAGACTCGAACAACAGTTCCTGACACAGATGATGGCCATTTGAA 2217  
QY 1849 GAAGTGAGCTTCCGATTTAACCAAGACCTGGAATTTGATATCTTCTTCTGCTCCAG 1908  
DB GAAATGGGAGATGTGTATGAAACAGACATGAGAGCTTTAGTGTGCACTTCTGATTTAA 2277  
QY 1909 CACTGTCCGATTTTGGGAAATTTGGGATGATCAAGAGGATCTTCCAGAGATGAG 1968  
DB TTCAAGCGCACGGAAGAGCTTCACTGATTTGAGGGGACAGG-----CAGCACATATC 2310  
QY 1969 TCCTGAGAGATCTCTGTGTGCTCTTATGATGAGGGATTAAGACCTCATTTAGAGAG 2028  
DB AACATGAGGCTCCACATGATGCTGTTCAGGTGGG---TCCAGTCAAGATGCC 2385  
QY 2029 CAGTGGAGAAATTTCTGCTCATGCTTGGCAACCAACCACTGCGGAGCTGAGCTG 2088  
DB TATTGGAGATTTCTTCTCTGCTCTCAAGGTACCAAGAACTGAAAGAGCTGACTA 2445  
QY 2089 GGCAGACATCTTGACAGAGCGGGCCATGAAAGCCTGTGTGCAAGCTGAGGCAATCCC 2148  
DB AGTGAATCTGCGAGCCACTGTGACATGAGAGGCTTTGTAAGACCTTGAAGACGCCCT 2505  
QY 2149 ACCTGCAAGATACGACCTCGATGTT---TAGAAATGACAGATTAACCTCGTGTGAG 2205  
DB CGTGCCTCTGAGAGACCTCGGCTGTGGCTGTGGCTCTCAAGCTGAGGACTGCAAG 2565  
QY 2206 CACTTGGAGAAATCTGATGACCAACCTGTAACCTTAAGATCCCTCAACTTGGAGGACC 2265  
DB GACTTGGAGGCTGAGAGCCCAACCAACCTGACCGAGCTGAGACTGAGCTTCAAT 2625  
QY 2266 CACTTGAAGAGAGATGTAAGAGTGGCTGTGAAGCTTTAAACACCCAAATGTTTG 2325  
DB GTGCTACAGATGTGAGGCAACACCTTTGCAAGAGACTGAGACGCGAGCTGCAAG 2685  
QY 2326 TTGAGATCTTTGAGGCTGATGTTGTGATGATGACCATGCTGTTAACCTGAAAGATCTCC 2385  
DB CTACAGGAGCTGAGGCTGTGATGCTGTGAGCTTCAAGTGTCTGCTGAGGACTGAGC 2745  
QY 2386 CAATCTTACGACCTCCCGAGGCTGAAATCTTGAAGCTGCGAGAGAAACAGGTGACA 2445  
DB TCTGTGTTAGTGCACGCCAGGCTGAAGAGATTAACCTTGCAGCAGAAACACTGGAT 2805  
QY 2446 GACCAAGAGATGCTCTTCAATGATGCTTGG 2480  
DB GACGTTGGCGGCGACTGCTCTGAGAGGGGCTCAG 2840

RESULT 4  
US-09-388-221B-5  
; Sequence 5, Application US/09388221B

Patent No. 6818750  
GENERAL INFORMATION:  
APPLICANT: Reed, John C. 6818750e1 Card Proteins Involved in Cell Death Regulation  
TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation  
FILE REFERENCE: P-LJ 3650  
CURRENT APPLICATION NUMBER: US/09/388,221B  
CURRENT FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 5  
LENGTH: 4332  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(4332)  
US-09-388-221B-5

Query Match 5.5%; Score 220.6; DB 4; Length 4332;  
Best Local Similarity 48.0%; Pred. No. 1,66-54;  
Matches 939; Conservative 0; Mismatches 944; Indels 72; Gaps 8;

Qy 532 TTTGATTCAGACCGGTGGGCTTCCGCTCGACAGTGTCTGTCAGGAAAGTCAGAA 591  
Db 952 TTTGGCCAGGCTTGATATCCAGAACCTGCGCATATGATATGTCAGGGGCTGCTGA 1011  
Qy 592 ATTGGAAATGGCTCTAGCCAGAAAGATCGTGTGCTGGGCGCAAGTGAATCTAC 651  
Db 1012 ATTGGAAATGCAACCTGGCCAGGCGGTGAAGAGCTGGGGGAGAGGCGCATAT 1071  
Qy 652 CAGGAAATGTTCTCTACGTTCTTCCCTCCCGTTAGAGATGACAGGGAAGAG 711  
Db 1072 GGGAGCCGCTTCAGCATGTTCTTCTTACCTGACAGAGCTGAGCCAGTCCAGAGT 1131  
Qy 712 AGCATGTTCAGAGATTCATCTCAGAGAGTGGCCAGACTCCAGAGCTCCGCTGACGAG 771  
Db 1132 GTAGTCTCGTGAATCATGGGAAAGATGGGACAGCACTCGGCTCCATTAGACAG 1191  
Qy 772 ATCATGTCCGACCAAGAAAGCTGTGTTCATCATTTAGACGTTTGAATGACTGGGCTCT 831  
Db 1192 ATCTGTCTAGGCGCAGAGCGGCTGCTCTTCAATCTCGATGTGTAGATGAGCAGAGATGG 1251  
Qy 832 GTTCTCAACATGACA---CAAGCTCTGCAAGAGCTGGGCGAAGACAGCTCCGCTTC 888  
Db 1252 GTCTTGACAGAGCCAGTTCGATCTGTCTGTCAGCTGAGCCAGCCACAGCCGGAGAT 1311  
Qy 889 ACCCTCATAGCAGCTGCTGAGAAAGTCTGCTCCGAGTCTTCCTGATCTGAC 948  
Db 1312 GCATGCTGGGAGTTTCTGGGAAACTATATCTCCGAGGCACTCTTCTGATCAAG 1371  
Qy 949 GTCAAGAGCTGGGACAGAAAGTCAAGTCAAGGTCTGCTCCCGTTACCTGTTA 1008  
Db 1372 GCTGGAACACAGCTCTGCAAACTCATTCCTTTGGAGCAGGACAGTGGGTAGAG 1431  
Qy 1009 GTTAGAGAAATCTCCGGGAAACAAGATCCATCTGCTCTTGGAGCGGGATTTGTAG 1068  
Db 1432 GTCTGGGATCTCTGATCTCAGCAGGAAAGAAATTTCTACAGATTTTTCACAGATGAA 1491  
Qy 1069 CATCAGAGACACAGAGGTGTGCTGATCAATGAACACCGTGAAGCTGTCAGCAGTGC 1128  
Db 1492 AGGCAAGCAATTAGAGCTTTTAGTGGTCAATCAAAACAAGACTCTGGGCTCTGTGT 1551  
Qy 1129 CAGGTGCGCGCGTGTCTCTCATCTGCTGGGCTCTGACAGTTCAGAGCGTGTGGG 1188  
Db 1552 CTGTGGCTGGGTCTCTGCTGGCTGACATTCCTGATGACAGCAAGTGAAGGGAAG 1611  
Qy 1189 GAGAGCGTGGCGCTTCAACCAACGCTCAAGGCTGACAGCGGCTTTGTGTAT 1248  
Db 1612 GAAAAACTCACATGACTTCAGAGCAACCAACACCTCTGTCTACATTAACCTTGCCAG 1671  
Qy 1249 CAGCTACCCCTCGAGGCGTGTGCTGAGGCTGTCTCAATCTGAGGAAAGATTTCTG 1308  
Db 1672 GCTCTCAAGCTCAGCAATGG-----GACCCAGCTC 1704

Qy 1309 AAGGCTTCTGCGGATGCTGTGGAGGAGGTGAATAGAAAGTCAGTGTGATGCT 1368  
Db 1705 AGAGACTCTGTCTCTGGCTGTGAGGAGCATCTGGCAAAAGAACCTTTTACGTCA 1764  
Qy 1369 GACGACTCATGTTTCAAGAGCTGGGAGTCTGAGCTCCGTCTCTGTTCAATGAC 1428  
Db 1765 GATGACCTCAGAGAGCATGGGTATGATGGGCAATCATCTCACCTCTTGAAGATGGT 1824  
Qy 1429 ATCTTCTCCAGACAGGCACTGAGAGTACTACCTTTCTTCACTCAGTCTCAG 1488  
Db 1825 A---TTCTTCAAGAGCAACCCATCCCTGAGCTTACAGCTTCACTCTGTGTCCA 1881  
Qy 1489 GACTTCTGCGCGCTTGTACTAGTGTGTAGAGGCGCTGGAATAGACCGAGCTCTGC 1548  
Db 1882 GAGTCTTGTGACGAATGTCTATGTCTTGAAGATGAGAGGAGGAGTAAATTTCT 1941  
Qy 1549 CCTCTGACGTTGAGAAACAAAGAGGTCATGAGCTTAAACAGCAGGCTTCATATC 1608  
Db 1942 AATTGATCATATGATTTGAAAGAGCCTAGAGATATGAAATATGATGCTGTTGG 2001  
Qy 1609 CACTGCTTTGATGAGAACGTTTCTTGTGGCTCTGAGCGAAGACGTAAAGAGCCA 1668  
Db 2002 GCATC-----AACCAACAGTTTCTTATTTGGCTGTTAATGATGAGGGGAGAGAG 2055  
Qy 1669 CTGAGATCTCTGCTGGCTCTCCGTTCCCTGGGGGGTGAAGCAGAACTTCTGACTGG 1728  
Db 2056 ATGAGAAACATTTTCACTGCGGCTGTCTAGGGGAGAAACCTGATCAGTGGTCCG 2115  
Qy 1729 GTCTCTGTGTGGGTGACAGCACTTAATGCAACACCCAGAGAGACCTCTGAGCCTTC 1788  
Db 2116 TCCCTGAGCTGTGCTGACAC-----ACACTCTGTGAAGTCTTC 2157  
Qy 1789 CACTGCTTTTGAAGACTCAAGACAAAGATTTGTGCTTGGCATTAACAGCTTCCA 1848  
Db 2158 CACTGCTTTGACGAGACTCGAACAACAGTTCTTCAACAGAGATGAGCCATTTGCA 2217  
Qy 1849 GAAATGTGCTTCCGATTAACAGAACTGAGATTTGATGATCTTCTTCTGCTCCAG 1908  
Db 2218 GAAATGGGCAATGTGTGAAGAAACAGACATGAGAGCTTTAGTGTGATCTTCTGCAATTA 2277  
Qy 1909 CACTGCTGATTTGGGAAATTCGGGTGATGTCMAAGGATCTTCCAGAGATGAG 1968  
Db 2278 TTCAAGCCGACGTAAGAGACTTCACTGATTTAGAGGAG-----CAGCAGATC 2330  
Qy 1969 TCCGCTGAGGATGCTGTGCTCTCTATGATGTCGGGATTAAGACCTCATTTAGAG 2028  
Db 2331 AACTGAGAGCCCAACAGATGATGCTGTTCAAGTGG-----TCCAGTCAAGATGCC 2385  
Qy 2029 CAGTGGAGAAATTTCTGCTCATGCTTGGACCAACCACTGCGGAGCTGAGCTG 2088  
Db 2386 TATTGGAGATTTCTTCTCCGTCTCAAGGTCAACAGAAACCTGAGAGACTGAGACTTA 2445  
Qy 2089 GGAGAGCATCTGACAGAGGAGCATGAAGACCTGTGTGCAAGCTGAGGATGCC 2148  
Db 2446 AGTGAAGCTGCTGAGCACTTGAAGTGAAGTCTTTTGAAGACCTTGAAGAGCTCT 2505  
Qy 2149 ACTGCAAGATACAGACCTGATGTT---TAGAATGACAGATTAACCTCTGTGTGAG 2205  
Db 2506 CGCTGCTCTGAGAGACCTGCGTGTGGCTGTGCTGCTCAAGCTGAGAGACTGAG 2565  
Qy 2206 CACTCTGAGAAATGCTATGAGCAACCTTAACCTTAAGATCTTCAACTTGGAGGAC 2265  
Db 2566 GACCTGCTTTTGGAGTGAAGCAACAGACCTTGAACCGAGCTGAGCTTGAAGTCAAT 2625  
Qy 2266 CACTGAGAGAGAGATGTAAGATGGCGTGTGAAGCTTAAACCCCAAAATGTTG 2325  
Db 2626 GTGCTACAGATGCTGAGCCAAACACTTTGCAAGACTGAGACAGCCAGCTGCAAG 2685  
Qy 2326 TTGAGTCTTTGAGGCTGATTTGCTGATGATGACCAATGCTGTTACCTGAAGATCTC 2385  
Db 2686 CTACAGGAGACTGAGCTGTGCTGAGCTGTGCTCAAGTGTGCTGTGCTGCAAGACTGAG 2745



Query 2386 CAAATCTTACGACCTCCCGAGCTGAAATCTCTGAGCCTGCGAGGAAACAAGGTGACA 2445  
Db 2746 TCTGTGCTTAGTGCAGCCCGAGCTTGAAGAGGCTAGACCTGCGAGCAACAACCTTGAT 2805  
Qy 2446 GACCAAGGAGTAAATGCTCTCTCACTGATGCTCTTGA 2480  
Db 2806 GACGTTGGCGTGCAGCTGCTCTGTGAGGGGCTCAG 2840

RESULT 5  
US-09-388-221B-11  
; Sequence 11, Application US/09388221B  
; Patent No. 6818750  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; TITLE OF INVENTION: No. 6818750el Card Proteins Involved in Cell Death Regulation  
; FILE REFERENCE: P-IJ 3650  
; CURRENT APPLICATION NUMBER: US/09/388,221B  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 4466  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(4272)  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Construct  
US-09-388-221B-11

Query Match 5.5%; Score 220.6; DB 4; Length 4466;  
Best Local Similarity 48.0%; Pred. No. 1.6e-54;  
Matches 939; Conservative 0; Mismatches 944; Indels 72; Gaps 8;

Qy 532 TTTGATTCAGACCGGTGGGGCTTCGCGCTCGACAGGTGTTCTGACGAAAGTCAGGA 591  
Db 952 TTTGGCCAGGCGTGGATACCCAGAAACCTGCGATAGTCATAGTCAGAGGGGCTGTGGA 1011  
Qy 592 ATTGGAAATCGGCTCTAGCCAGAAAGATCGTCTGTGCTGCGGCGCAAGTGAACCTTAC 651  
Db 1012 ATTGGGAATGCAACACTGGCCAGAGGTGAAGAGCCTGGGGAGAGGCGCAGCTGTAT 1071  
Qy 652 CAGGAATGTTCTCTACGTCT 711  
Db 1072 GGGGACCGCTTCAGCATGCT 1131  
Qy 712 AGCAGTTCACAGATTCATCTCCAGAGGTGCGCAGACTCCAGGCTCCGATGACGAG 771  
Db 1132 GTGAGTCTCGTGAAGTCTATGGAAGATGGAGCAGCCTCCGCTCCCATTAAGACG 1191  
Qy 772 ATCATGTCGACCAAGAAAGCTGTGTTTCAATTAAGGTTTCATGACCTTGGGCTCT 831  
Db 1192 ATCTGCTTAGGCGACAGCGGCTGCTCTTCACTCCGATGTGTATGATGACCAAGATGG 1251  
Qy 832 GTCTCAACATGACA---CAAAGCTTCGAAAGCTGGGCTGAGAAAGCAAGCTCCCTTC 888  
Db 1252 GTCTTGCAGAGCCGAGTTCGAGCTGTCTGTGACGTGAGCCAGCCACACCGCGGAT 1311  
Qy 889 ACCCTATACGAGTCTGCTGAGAGGTCTGCTGCTCCCTGAGTCTCTTCTGATCGTCAAC 948  
Db 1312 GCACTGCTGGGCAATTTCTCTGGGAAACTATATCTCCGAGGATCTCTTCTGATACG 1371  
Qy 949 GTCAAGAGCTGGGCAAGAGAGTCAAGTCAAGAGTGTCTCTCCGTTACTCTGTTA 1008  
Db 1372 GCTGCGACCAAGCTCTGCAAGACCTCACTCTTCTTGGAGCAGGCACTTGGGTTAG 1431  
Qy 1009 GTTAGAGGATCTCCGGGGAAACAAGATTCATTTGCTCTCTGAGCGGGGATTTGAG 1068  
Db 1432 GTCTGGGGTCTCTGATCCAGAGGAAAGATTTTCTACAGATATTTTCAAGATGAA 1491

Qy 1069 CATCAGAAACACAGGGTTGCGTGCATGTAACAACCGTGAAGTGTCTGACCACTGTC 1128  
Db 1492 AGGCAAGCAATTAGAGCTTTAGGTTGTCAATTAACAAGAGCTCTGGGCTCTGTGT 1551  
Qy 1129 CAGGTCCCGCGTGGGCTCTCTCATTTGCGTGCCCTGCAAGCTGCGAGAGTGTGGG 1188  
Db 1552 CTGTGTCCCTGGGTTGTCTGGCTGGCTGCACTTGTGATGCGAGAGTGAACGGAG 1611  
Qy 1189 GAGAGCGTGGCCCTTCAACCAAGCTCACAGGCTGCAAGCGCGCTTTTGTGTTTCAT 1248  
Db 1612 GAAAACTCACACTGACTTCCAGAACCAACCAACCTCTGTCTACTATACCTTGCCAG 1671  
Qy 1249 CAGCTCAACCCCTCGAGGCGTGTCCGCGCTGTCTCAATCTGAGAAAGTTGCTCTG 1308  
Db 1672 GCTCTCAAGCTCAGCCATTGG-----GAGCCCACTTC 1704  
Qy 1309 AAGGCTTTCGCGTATGCTGTGAGAGGAGTGTGAATAGAAAGTCAAGTCTTGAATGT 1368  
Db 1705 AGAGACCTGTGCTCTCGGCTGTGAGGGCATCTGGCAAAAAGAACCTTTTCAATCCA 1764  
Qy 1369 GAGAGCTCAATGTTCAAGAGCTGGGAGTGTGAGCTCCGCTGCTGTTTCAATGATAC 1428  
Db 1765 GATGACCTCAGAAAGCATGGGTTAGATGGGCCATCATCTCCACTTTTGAAGATGGT 1824  
Qy 1429 ATCTCTTCCAGACAGCCCACTGTGAGAGTACTACACTTCTTCAACCTTCACTGTCAG 1488  
Db 1825 A---TTCTTAAGAGCAACCCCATCCCTGTAGCATAGAGCTTCACTTCTGTTCAA 1881  
Qy 1489 GACTTCTGTGCGGCTTGTACTAGTGTGTAGAGGCTGGAATTCAGCCAGCTCTGTC 1548  
Db 1882 GAGTCTTTGAGGAAGTCTTATGTCTTGAGAGTGAAGAGGAGGATTAACATTTCT 1941  
Qy 1549 CCTCTAGCTTGAAGAACAAAGAGTCCATGAGCTTAAACAGCAGGCTTCATATTC 1608  
Db 1942 AATTGCACTATAGATTTTGAAGAAACCTTAGAAGCATATGAAATACATGAGCTTTGGG 2001  
Qy 1609 CACTGCTTTGATGAAGAGCTTTCTGTTTGGCTGTGAGCGAAGACGTAAGAGAGCCA 1668  
Db 2002 GCATC-----AACCAACGTTTCTTATGTGGGCTGTAAAGTATGAGGGGAGAGAG 2055  
Qy 1669 CTGAGGCTCTGCTGGGCTGTCCGTTCCCTTGGGGGTGAAGCAAGCTTCTGACTGG 1728  
Db 2056 ATGGAAGAACTTTTCACTCGCGCTGTCTCAGGGGAGGAACCTGATGCAAGTGGTCCG 2115  
Qy 1729 GTCTCTGTTGGGTGACAGCCTAATGACCAACCCAGAGACACCTCGAGAGCCTTC 1788  
Db 2116 TCCCTGAGCTGCTGCTGAGCC-----ACACTCTTGGAGTCCCTC 2157  
Qy 1789 CACTGTCTTTTGAAGCTCAAGACAAAGAGTTTGTGCTTGGCATTAACAGCTTCCA 1848  
Db 2158 CACTGCTTTGAAGAGCTCGGAAACAAACGTTCTGACAAAGATGATGGCCATTTCGA 2217  
Qy 1849 GAAGTGGCTTCCGATTAACAGAACCTGATCTTGAATAGCATTTCTTCTGCTCAG 1908  
Db 2218 GAAATGGGCAATGTGTGAAGAACAGACATGAGCTCTTAAGTGTGCACTTTCTGATTA 2277  
Qy 1909 CACTGTCCGATTTTGGGAAATTTGGGTGGAATGCAAAAGGATCTTCCAAGATGAG 1968  
Db 2278 TTCAAGCCGCACTGAAGAAAGCTTCACTGATTAAGGACAG-----CAGCACAATC 2330  
Qy 1969 TCCGCTAGGAGATGCTCTGTGCTCTCTATGATGCGGATTAAGACCTCATATTGAGAG 2028  
Db 2331 AAGATGAGGCCCAACATGATGATGCTGTTCAGGTGG-----TCCAGTCAAGATGCC 2385  
Qy 2029 CAGTGGAAATTTTCTGCTCATGCTTGGGACCAACCAACCTGCGGCACTGAGCTTG 2088  
Db 2386 TATTTGGAGATTTCTCTTCTCCGTCCTCAAGTCAACGAAACCTGAAAGAGCTGAGCTTA 2445  
Qy 2089 GGAGAGCACTCTGACAGAGCGGGCCATGAAAGACCTGTGTGCAAGCTGAGGATCCC 2148  
Db 2446 AGTGAAGTCTGCTGAGCCTCTGCAAGTGAAGATTTTGTGAAGACCTTGAAGCGGCT 2505  
Qy 2149 ACTGCAAGATACAGACCTGATGTT---TAGAAATGACAGATTAACCTCTGTGTGAG 2205

Db 2506 CGCTGCTCTCTGAGACCTCTGGGGTTGGCTGGCTGGCTCAGAGCTAGAGCTGCAAG 2565  
Qy 2206 CACCTCTGAGATGTCATGCGCAACCTTAACCTTAAGTCCCTCACTTGGAGGCAACC 2265  
Db 2566 GACCTTGGCTTGGCTGAGACCAACCTTAACCTTAAGTCCCTCACTTGGAGGCAACC 2265  
Qy 2266 CACCTGAGAGAGAGATGTAAGATGAGCTGTAAGCTTAAACACCAAAATGTTG 2325  
Db 2626 GTGCTCAGAGATGCTGAGACCAACCTTGGCAGAGCTGAGAGAGCTGCAAG 2685  
Qy 2326 TTGAGCTCTTGGAGCTGATGCTGTAAGTGAACCTGTTAACCCTGAAGATCTCC 2385  
Db 2686 CTACAGCAGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAG 2745  
Qy 2386 CAATCTTGAACCTCTCCCGAGCTGTAAGTCTGAGCTGAGCTGAGCAAGGATGACA 2445  
Db 2746 TCTGTGCTTGAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2805  
Qy 2446 GACGAGGAGTAACTGCTGAGTATGCTGAG 2480  
Db 2806 GACCTTGGCTGCTGAGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2840

## RESULT 6

US-08-910-731-1

Sequence 1, Application US/08910731

Patent No. 5932440

GENERAL INFORMATION:

APPLICANT: CHATTERJEE, DEB K.

APPLICANT: SHANILYA, HARINI

TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESS: STERN, KESSLER, GOLDSTEIN &amp; FOX P.L.L.C.

STREET: 1100 NEW YORK AVE., N.W., SUITE 600

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/910,731

FILING DATE: (Herewith)

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/795,395

FILING DATE: 04-FEB-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/794,546

FILING DATE: 03-FEB-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/024,057

FILING DATE: 16-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: ESMOND, ROBERT W.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0942,3440003

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1371 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS  
LOCATION: 1..1368  
US-08-910-731-1

Query Match 4.6%; Score 183.8; DB 2; Length 1371;  
Best Local Similarity 47.2%; Pred. No. 7e-44;  
Matches 591; Conservative 0; Mismatches 657; Indels 3; Gaps 1;

Qy 2016 CCTATTAGAGAGCAGTGGAGAAATTTCTGCTCCATGCTTGGACACCCACACCTGCG 2075  
Db 102 CCTACGAGAGAGCAGTGGAGAAATTTCTGCTCCATGCTTGGACACCCACACCTGCG 161  
Qy 2076 GAGCTGAGACCTGGGAGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2135  
Db 162 GAGCTGAGACCTGGGAGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 221  
Qy 2136 GCTGAGAGATCCACCTGAGAGATCAGACCTGAGAGAGAGAGAGAGAGAGAGAGAG 2194  
Db 222 CTTGAG 281  
Qy 2195 --CTGCTGAG 2252  
Db 282 GAGCTGAG 341  
Qy 2253 CTTGAG 2312  
Db 342 TCTCAG 401  
Qy 2313 CCCAAATGTTTGTGAG 2372  
Db 402 CCCCCAG 461  
Qy 2373 CTTGAGAGATCTCCAAATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2432  
Db 462 GAG 521  
Qy 2433 AAG 2492  
Db 522 CAG 581  
Qy 2493 CGCCCTGAG 2552  
Db 582 CAG 641  
Qy 2553 GAGCTGAG 2612  
Db 642 GTGAG 701  
Qy 2613 GAG 2672  
Db 702 GAG 761  
Qy 2673 GAG 2732  
Db 762 GAG 821  
Qy 2733 GCTTATGAG 2792  
Db 822 CTTGAG 881  
Qy 2793 TGGAG 2852  
Db 882 GAG 941  
Qy 2853 GTTGTATGAG 2912  
Db 942 GGTATGAG 1001  
Qy 2913 GAG 2972  
Db 1002 GAG 1061  
Qy 2973 TGGAG 3032





Db 1301 TGACGACATTTACTGTCTGTGAGGATGAGAGACCGGCTGCGACCCCTGGAGAAAGACA 1360  
QY 3257 AGCC 3260  
Db 1361 AGCC 1364

## RESULT 9

US-08-910-731-7

Sequence 7, Application US/08910731

Patent No. 5932440

GENERAL INFORMATION:

APPLICANT: CHATTERJEE, DEB K.

APPLICANT: SHANDILYA, HARINI

TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN &amp; FOX P.L.L.C.

STREET: 1100 NEW YORK AVE., N.W., SUITE 600

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/910,731

FILING DATE: (Herewith)

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/795,395

FILING DATE: 04-FEB-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/794,546

FILING DATE: 03-FEB-1997

APPLICATION NUMBER: 60/024,057

FILING DATE: 16-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: ESMOND, ROBERT W.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0942.3440003

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2540

TELEPHONE: 202-371-2600

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1371 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

MOLECULE TYPE: cDNA

US-08-910-731-7

Query Match 4.5%; Score 180.8; DB 2; Length 1371;

Best Local Similarity 47.1%; Pred. No. 5.5e-43;

Matches 588; Conservative 0; Mismatches 657; Indels 3; Gaps 1;

QY 2016 CCTCATGAGAGGAGTGGAGAAATTTGCTCCATGCTGAGGACCCACCCACCTGAGG 2075  
Db 102 CCTCAGGAGGAGGACCTGCAAGACATGCTGCTCCCTCCGAGGCAACCTCCCTGAC 161  
QY 2076 GCAGCTGACCTGCGACGAGCATCTGACAGAGCGGGCCATGAAACCTGTGTGCCAA 2135  
Db 162 CGAGCTGCGCTCGGACCAAGAGCTGCGGCGATGCCGCGTGACCTGTGTGCAAGG 221  
QY 2136 GCTGAGGATCCCATCTGCAAGATACAGACCTGTATGTTTGAAGAAATGACAGATTACC- 2194  
Db 222 CCTGCAAGAGCCCACTGCAAGATCCAGAAAGCTCAGCTCGAGAACTGCTCCCTGACCGA 281

QY 2195 --CTGATGTGACGACCTCTGAGAAATGCTATGCGCCACCGTAACTTAAGATCCCTCAA 2252  
Db 282 GCGCGGCTGCGGGGTCTCTGCGCAGACAGCTGCGCTCCCTGCCACGCTGGGAGCTGCA 341  
QY 2253 CTGGAGAGGACCCACCTGAAAGAAAGATGTAAAGATGCGGTGTAAGCTTTAAAAACA 2312  
Db 342 TCTCAGCGACAAACCACTGCGGGAGACCGGCGCTGCGCTGCTGTGTAGAGGGCTCTGGA 401  
QY 2313 CCGAAATGTTTGTGAGAGCTTTGAGGCTGAGATTGCTGTGAGATTGACCATGCTGTGTA 2372  
Db 402 CCCCAGTGCACCTGGAAGAGCTGCAATTGAGATGCTGCGCTGACCGCGCCACCTG 461  
QY 2373 CCTGAAGATCTCCCAATTCCTTAAGACCTGCGCCCGACCTGAATCTGTGAGCTGGCAGG 2432  
Db 462 GCAAGCCCTGCGCTCGGTGCTCAGGGGCAACCGGCGCTTGAAGAGCTCAGGTGACCA 521  
QY 2433 AAACAAGGTGACAGACCAAGGAGTAAATGCTTCAAGTATGCTTGAAGAGTCTCCAGTG 2492  
Db 522 CAACGACATCGGCGAGGCGCGGCGCTGCGGCGCAAGGATGCGGCGGCTGCGCTG 581  
QY 2493 GCGCTGCAAGAGCTGATCTGAGAGCTGTGAGCTATACAGCCACGCGTTGCCAGAGCT 2552  
Db 582 CCAAGTGAAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 641  
QY 2553 GAGCTGAGCCCTGCTGAGCAACCGGAGCTTGAACACCTGAGCTTATCCAAACAGAGCT 2612  
Db 642 GTGCGGAATTTGTGCTCTCCAGGCTGCTGCGGAGCTGCGGCTGCGGAGCAAGAGCT 701  
QY 2613 GGGGAACGAAGGTAAATCTACTGTGTGATTCATGAGGCTTCCCACTGTATGCTGCA 2672  
Db 702 GGTGATGTGGGCAATGAGCGAGCTGCGGAGGCTGCTCCAGCTTCCAGGCTCAG 761  
QY 2673 GAGGCTGATGCTGAATAGTGGCCACTGAGACAGGCTGCTGTGCTTTCTGTGACTGTC 2732  
Db 762 GACCTGTGATCTGGAAGTGTGCACTGCTCCAAAGGCTGCGGAGATCTGTGCTGCT 821  
QY 2733 GCTTATGAGTAACTCATGAGCTGAGCAGCACTGAGCTTATGAGATGAGTAAAGACAA 2792  
Db 822 CTTAGAGGCAAGAGAGAGCTGAAAGAGCTGAGCTGCGGCGGCAACAGCTGCGGAGTGA 881  
QY 2793 TGCGGTGAAGTTCTGTGCGAGGTGATGAGAAACATCTTGTGATCTGAGAGCTGGA 2852  
Db 882 GGTGCGCCAGCTGCTGTGTGAGACCTGTGTGAACTGCTGCGAGCTGAGTGTGCTGTG 941  
QY 2853 GTTGTAAGTGTATCTCACCGCGCTGTGTGAGAGTGTGCTGTGTGTGATCTGAG 2912  
Db 942 GGTGAAGTCTGCAAGCTTCAAGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1001  
QY 2913 GAGCAGACACTGAAGAGCTGATCTCAGGACCAATGCTGCGGTGAGCGGTGCGGTGCTG 2972  
Db 1002 GAACAGTTTCTCTGAGAGCTACAGATTAAGCAACAGAGCTGAGAGATGCGGCGTGG 1061  
QY 2973 TGCGCTGTGCGAGGAGCTGAAAGCAAAAGATGTTCTGACAGAGCTGCGGTGAAGGC 3032  
Db 1062 GGAAGTGTGCGAGGCGCTGCGGCGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1121  
QY 3033 ATGTGACTGACTTGTGATGCTGTGAGGACCTTCTGTGCGCTTCTGTGCAACCGGCA 3092  
Db 1122 CTGGAATGTGATGACAGAGCTGACAGAGCTGCGCAACCTGTGTGCGCAACGAG 1181  
QY 3093 TCTGACAGTAAACCTGTGAGAGTAACTTCAATGCTGCAAGAGATGATGAGCTGTG 3152  
Db 1182 CTGTGTGAGTGTGACCTTCAAGCAACAGCTGTGAGGAGACCGGCGATCTGTGAGCTGTG 1241  
QY 3153 TTGCGCTTGTGCTGCTGCGACAGTCTTAATTAAGATTAATGAGTGTGAAATGCGAGTA 3212  
Db 1242 GGAAGAGGTGCGGAGCGGCGCTGCTGTGAGACAGCTGATGCTGTGAGCAATTTACTG 1301  
QY 3213 CCTGTGCAATTAAGAAAGCTGTGAGAGAGTACACTGCAAGCC 3260  
Db 1302 GTTGAAGAGATGAGAGACCGGCTGAGGCGCTGGAAGAGCAAGGCC 1349



```

; Patent No. 5965399
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Cloning and Expression of Rat Liver and
; TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,395
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2540
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1374 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1368
; US-08-795-395-3

```

```

Query Match      4.4%; Score 175.8; DB 2; Length 1374;
Best Local Similarity 47.2%; Pred. No. 1.7e-41;
Matches 568; Conservative 0; Mismatches 632; Indels 3; Gaps 1;

```

```

QY 2000 GGATCGGGGATTAAGACCTCTCATTTGAGAGCAGTGGGAGGATTTCTGCTCCATGCTTGGCA 2059
DB 86 GGGTGAAGTACTGTGGCTCTCACTGAAGTGGGTGCAAGATCAATGATCAAGATCCAGG 145
QY 2060 CCCACCCACCTGCGGAGCTGAGCTGACCTGGGAGAGCATCTTGAACAGAGCGGCGCATGA 2119
DB 146 CCAACCTGACCTGACAGAGCTCAGCTACGACCAATGAATGATGATGCTGGTGG 205
QY 2120 AGACCTGTGTGCCAGGTGAGGATCCCACTGCAATACAGACCTTGATGTTTAA 2179
DB 206 GTCTGTCTCTCGAGGCTGAGGATCCCACTTGTGAATCCAGAGGCTGAGCCCTTCAGA 265
QY 2180 ATGCACATATTACCC--CTGGTGTGACAGACCTTGAGAAATGCTCAATGGCCAAACGTA 2236
DB 266 ACTGAGCTTGAAGGAAAGCTGCTGTGGGTCTCTCCGATGATGCTGGCTCTTTGTCTTA 325
QY 2237 ACTTAAGATCCCTCACTTGGAGGACACCACTTGAGAGAAAGAGATGAAGATGGCGT 2296
DB 326 CCTGCGTGAATCAATCATGATGACAAACCTCTGGGGAGTGAAGCTGAAGCTGCTCT 385
QY 2237 GTGAAGCTTAAACACCCAAATGTTTGTGATGATCTTTGAGGCTGATGCTGTGAT 2356
DB 386 GTGAAGGACTCGGGAGCCCAAGTCCGCTTTGAGAAAGCTTCAATTGAAATATGTAACC 445

```

```

QY 2357 TGAACCATGCTGTATTACCTGAGAGATCTCCAAATCCTTAGACCTCCCGAGCTGAAT 2416
DB 446 TCAAGCTACAGCTGAGAGCCCTGGCTCTCAGTGTCAAGGTGAACCTGACTTTAAAG 505
QY 2417 CTGAGCTGGCAGAGAAACAGGTGACAGACAGGAGATGATGCTCTCAATGATGCT 2476
DB 506 AGCTAGTATTGAGCAACAATGATCTTCATGAGGCTGATTCACACTCTGTGCCAGGACC 565
QY 2477 TGAAGTCTCCAGTGGGCCCTGAGAGGCTGATGAGGAGATGAGGATGAGGATCAAGCCA 2536
DB 566 TGAAGATTTCTGCTCTGCAACTGAGATCACTCAACTGAGAAATCTGTGATTCATCAG 625
QY 2537 CGGGTTCAGAGATCTGAGCTTCAAGCCCTGTCAAGCAACCGAGCTTGACACACTGTGCC 2596
DB 626 CCAACTGCAAGGATCTGTGATGATGTTGTGGCTCAAGGCTTCACTCAAGAACTGGAAGT 685
QY 2597 TATCCAAACACAGCTTGGGAGAGCAAGGTGAATCTTACTGTGTGATTCATGAGGCTTC 2656
DB 686 TGGCAGACCAACAACTGGGCAACAGGCAATTCAGCACTGTGCTCAGAGACTGTGCTTC 745
QY 2657 CCCACTGATGCTGAGAGGCTGATGCTGATCAAGTCCCACTGAGACAGGCTGGCTGTG 2716
DB 746 CCACTGAGGCTGAGACTCTGTGCTCTGGGCTGTGATGATCACTGACAGAAAGCTGCA 805
QY 2717 GTTTCTTTCGACTTGGCTTAAGGTAATCATAGCTGAGCACTGAGCTTATGATGA 2776
DB 806 AGGACTGTGCTGTGCTCTGAGCCAGAGAGCTGAGGAATCAAGCTTACTGCTGAGCA 865
QY 2777 ACCCTGGAAGACATGAGCGTGAAGCTTCTGTGCGAGTGTGATGAGAAACATCTTGTTC 2836
DB 866 ATGAGCTGAAGGATGAGGGTGCACCACTGTGTGTGAGAGCTGTGATGAGCTGTGCTGC 925
QY 2837 ATTCACAGAACCTGAGATGAGTAAAGTATCATCTCAACGCGGCTGTGAGAGTCTGT 2896
DB 926 AGCTGAGTCACTGTGGTAAAGCTGTAGCTTCAAGCTGCTCTTGTGCCACTTCT 985
QY 2897 CTTGTGTATCTCGAGGAGACACCACTGAAAGCTGATATTCACAGCAATGCTGCTGG 2956
DB 986 GCTGAGTGTGACCAAAAAGTCTCTGTTTGAATGGAATGAGAGCAACCGGCTGG 1045
QY 2957 GTGACGTTGGGCTTGTGCTGCTGTGCGAGGAGCTGAAGCAAAAGAACATGTTTGAAGA 3016
DB 1046 GAGACTGGGAGTGTGAGGCTTGTGCAAGGCCCTGGGCTATCCGAGACAGTGTGCTGG 1105
QY 3017 GACTCGGTTGAAGGACATGTGGAATGACTTCTGATTTGCTGAGAGGCACTGCTTGGCCC 3076
DB 1106 TGTCTTGTGGGAGACTGTGATGTGACAGACAGTGTGAGAGCTTGTGCCACTGTTC 1165
QY 3077 TTTCTGCAACCGGACATCTGACAGCTTAAACCTGTGAGAAATTAATCTGATGCCAAG 3136
DB 1166 TGTGCGCAACCGGACGCTTGAAGGAACTGGAACCTCAATTAACATGATGGGGAGACAAG 1225
QY 3137 GAATGATGAAGCTGTGTTGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3196
DB 1226 GTGTCTTAACTGTGTGAGAGCTTCAAAAGCCCACTGATCTTCAAGAGCTGTTC 1285
QY 3197 TGT 3199
DB 1286 TGT 1288

```

```

RESULT 12
US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHRIFFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner

```



STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-9300  
TELEFAX: (703) 683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: PTZPT-Fls  
US-08-232-463-14

Query Match 1.5%; Score 61; DB 1; Length 7218;  
Best Local Similarity 1.1%; Pred. No. 1,2e-06;

Matches 4; Conservative 225; Mismatches 130; Indels 0; Gaps 0;

QY 264 GGAGAGAGCAGCAGATGATGATAAATTTACAGAGCTATGAAACAAGAGTGCCAC 323  
DB 1421 RRR 1362  
QY 324 AGCAGCAGAGCAGAGAAACAAGAAATTTACAGAGCTATGAAACAAGAGTGCCAC 383  
DB 1361 RRR 1302  
QY 384 AGCAGAGCAGAGAAACAAGAGCAGAGTGACATGAGGACTCAAGAGTCACGTGAT 443  
DB 1301 RRR 1242  
QY 444 GACCAATTCGCTGAGAGAGATGATGCTGCTGTTTGAACAACGTGCTGACTG 503  
DB 1241 RRR 1182  
QY 504 GCCGCAATTCAGCAAGCTGCTGCTGCTTTGATTCAAGCCGCTGCGCTTCGCGCTCG 563  
DB 1181 RRR 1122  
QY 564 CACGCTGCTTGCACGAGAACTCAGAAATTCGCTGCTGCTGCTGCTGCTGCTGCTG 622  
DB 1121 RRR 1063

## RESULT 13

US-09-023-655-395  
; Sequence 395, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart

APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
TITLE OF INVENTION: EXPRESSION  
NUMBER OF SEQUENCES: 1508  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023,655  
FILING DATE: HEREMITH  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 395:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: THPLBL02  
CLONE: 153338  
US-09-023-655-395

Query Match 1.5%; Score 60; DB 4; Length 376;  
Best Local Similarity 55.6%; Pred. No. 3e-07;

Matches 134; Conservative 0; Mismatches 106; Indels 1; Gaps 1;

QY 2914 AGCAGACACCTGAGAGAGCTGATCTCAGCAGCAATGCCCTGGGTGACGGTGGTGGT 2973  
DB 35 ACCAGAGCTGCTGAGAGCTGAGACTGAGTGAACAAGCTGCTGCTGCTGCTGCTGCT 94  
QY 2974 GCGCTGTGCGAGGAGCTGAGCAAGCAAGAGAGAGTGTCTGACGAGACTCGGCTGAAGGCA 3033  
DB 95 CTTCTGTGTGTGGAGAGCACTGTGTGCAATCTGAAGAGCTGCTGCTGCTGCTGCTGCT 154  
QY 3034 TGTGAGCAGACTTGTGATGCTGTGAGGCACTCTCTTGGCCCTTCTCTGAAACGGCAT 3093  
DB 155 TGTGCTCAGATGAGATGTTGAGATCTTNCATGATGATGAGACACAGCCATTC 214  
QY 3094 CTGACAG-CTTAACCTGTGAGAGATTAACCTGAGTCCCAAGGAATGATGAGTGTG 3152  
DB 215 CTGACAGACTCTATGTGTGGGAGAAATGCTTGGAGACTCAGAGATCGCAATTATG 274  
QY 3153 T 3153  
DB 275 T 275

## RESULT 14

US-09-949-016-16600  
; Sequence 16600, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

```

1 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
2 TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
3 FILE REFERENCE: CL001307
4 CURRENT APPLICATION NUMBER: US/09/949,016
5 CURRENT FILING DATE: 2000-04-14
6 PRIOR APPLICATION NUMBER: 60/241,755
7 PRIOR FILING DATE: 2000-10-20
8 PRIOR APPLICATION NUMBER: 60/237,768
9 PRIOR FILING DATE: 2000-10-03
10 PRIOR APPLICATION NUMBER: 60/231,498
11 PRIOR FILING DATE: 2000-09-08
12 NUMBER OF SEQ ID NOS: 207012
13 SOFTWARE: FastSeq for Windows Version 4.0
14 SEQ ID NO: 16600
15 LENGTH: 298336
16 TYPE: DNA
17 ORGANISM: Human
18 FEATURE:
19 NAME/KEY: misc_feature
20 LOCATION: (1)...(298336)
21 OTHER INFORMATION: n = A,T,C or G
22 US-09-949-016-16600

```

Query Match	1.3%	Score 51.4	DB 4	Length 298336
Best Local Similarity	58.0%	Pred. No. 0.011		
Matches 91; Conservative	0;	Mismatches 66;	Indels 0;	Gaps 0;

Qy	3272	TGACGCTAGTTCGATTCCTTTTGATGAAGTGCACGACACAAATAAGACTTACTTTCCG	3332
Db	202167	TTAGAGAACTGCTGATTATTATTATTCATTCATGCACTACACAAGGTAGTTGGCATTTCCC	2022266

QY	3333	GCTCCCTGAAAGCCGGGCAATGCGCATGTGCTTGGTGTGGGGGATGAACCCAGAGCAGAA	3392
Db	202227	ATTCTCTATACATTAAAGCATTAGCCATGTAACTTGTCTGTGCAATGAATCATGAAAGAA	202286

QY	3393	GAAGCGTGTGCGCTTCTGGCTGGAGACTTCAAGAGC	3429
Db	202287	GATATATATGTCACTCTGGGAGGAGCTTTAAGAGC	202323

RESULT 15  
US-09-949

```

US-09-949-016-14581/c
; Sequence 14581, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001107
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14581
; LENGTH: 64309
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(64309)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14581

```

Query Match	1.3%	Score 50.8;	DB 4;	Length 64309;
Best Local Similarity	58.7%;	Pred. No. 0.0059;		
Matches 88;	Conservative 0;	Mismatches 62;	Indels 0;	Gaps 0;

<b>Qy</b>	264	GGAGAAAGGCACGGGATGACATGAAAAAATTTCACAAGCTATGGAACAAAGAAGTGCCAC	323
<b>Db</b>	24025	GGAAGAGGAAAGAGAAAGAACGACMAAMAGAAAAGAGAGAAAGAAAGAAAGAA	23966
<b>Qy</b>	324	AGCAGCAGAGACAGAAAGAAACAAGAAATTTCAACAAGTATGSAACAAGAAAGTGTCCACGC	383
<b>Db</b>	23965	AGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA	23906
<b>Qy</b>	384	AGCAGAGACCAAGAAACAAGCATGTAGAG	413
<b>Db</b>	23905	AGAAAGAGAGAAAGAAAGAAAGAAAGAGG	23876

Search completed: July 19, 2005, 01:47:59  
Job time : 645 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2005, 16:47:53 ; Search time 2336 Seconds  
(without alignments)  
10958.514 Million cell updates/sec

Title: US-10-066-521-5

Perfect score: 4035

Sequence: 1 atggaagagagaacaatcgct.....tcgaggsgctgtgtctctaa 4035

Scoring table:

IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7173243 seqs, 3172129809 residues

Total number of hits satisfying chosen parameters: 14346486

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq:\*
- 20: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq:\*
- 21: /cgn2\_6/ptodata/2/pubpna/US10I\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*
- 23: /cgn2\_6/ptodata/2/pubpna/US11A\_NEW\_PUB.seq:\*
- 24: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 25: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*
- 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4035	100.0	4035	US-10-124-498-5	Sequence 5, Appli
2	4035	100.0	4035	US-10-066-521-5	Sequence 5, Appli
3	3166.2	78.5	3885	US-10-860-761-3	Sequence 3, Appli
4	3166.2	78.5	3900	US-10-399-443-23	Sequence 23, Appli
5	3166.2	78.5	3900	US-10-677-943-23	Sequence 23, Appli
6	3139	77.8	3830	US-10-216-645-3	Sequence 3, Appli
7	3133.6	77.7	3489	US-10-416-642-3	Sequence 3, Appli

8	3133.6	77.7	3926	15	US-10-216-645-1	Sequence 1, Appli
9	2900.6	71.9	3226	18	US-10-092-900A-347	Sequence 347, App
10	1278	31.7	3447	17	US-10-399-443-5	Sequence 5, Appli
11	1278	31.7	3447	18	US-10-677-943-5	Sequence 5, Appli
12	1152.8	28.6	1157	17	US-10-399-443-1	Sequence 1, Appli
13	1152.8	28.6	1157	18	US-10-677-943-1	Sequence 1, Appli
14	756.6	18.8	1075	17	US-10-399-443-3	Sequence 3, Appli
15	756.6	18.8	1075	18	US-10-677-943-3	Sequence 3, Appli
16	729.4	18.1	3521	20	US-10-794-342-2	Sequence 2, Appli
17	438	10.9	2775	17	US-10-357-820-51	Sequence 51, Appli
18	424	10.5	2985	10	US-09-965-621-15	Sequence 15, Appli
19	424	10.5	2985	17	US-10-407-866-15	Sequence 15, Appli
20	424	10.5	2985	19	US-10-781-294-15	Sequence 15, Appli
21	422.4	10.5	2982	14	US-10-124-498-25	Sequence 25, Appli
22	422.4	10.5	2982	14	US-10-066-521-25	Sequence 25, Appli
23	422.4	10.5	3190	17	US-10-054-749-1079	Sequence 1079, Ap
24	422.4	10.5	3368	14	US-10-124-498-23	Sequence 23, Appli
25	422.4	10.5	3368	14	US-10-066-521-23	Sequence 23, Appli
26	419.2	10.4	2575	9	US-09-848-035-7	Sequence 7, Appli
27	419.2	10.4	2575	9	US-09-986-224-7	Sequence 7, Appli
28	419.2	10.4	2883	17	US-10-161-493-3	Sequence 3, Appli
29	419.2	10.4	3172	19	US-10-398-037-74	Sequence 74, Appli
30	409	10.1	3186	14	US-10-124-498-17	Sequence 17, Appli
31	409	10.1	3186	14	US-10-066-521-17	Sequence 17, Appli
32	404.4	10.0	3108	10	US-09-965-621-23	Sequence 23, Appli
33	404.4	10.0	3108	17	US-10-407-866-23	Sequence 23, Appli
34	404.4	10.0	3108	19	US-10-781-294-23	Sequence 23, Appli
35	404.4	10.0	3218	17	US-10-407-866-67	Sequence 67, Appli
36	392.6	9.7	3189	15	US-10-239-653-10	Sequence 10, Appli
37	392.6	9.7	4944	22	US-10-499-353A-429	Sequence 429, App
38	390.6	9.7	3540	20	US-10-794-342-7	Sequence 7, Appli
39	381.4	9.5	3099	13	US-10-127-516-3	Sequence 3, Appli
40	381.4	9.5	3099	13	US-10-027-629-3	Sequence 3, Appli
41	381.4	9.5	3099	16	US-10-132-567-3	Sequence 3, Appli
42	381.4	9.5	3431	13	US-10-127-516-1	Sequence 1, Appli
43	381.4	9.5	3431	13	US-10-027-629-1	Sequence 1, Appli
44	381.4	9.5	3431	16	US-10-132-567-1	Sequence 1, Appli
45	368.4	9.1	2763	19	US-10-250-615-59	Sequence 59, Appli

#### ALIGNMENTS

RESULT 1  
US-10-124-498-5  
; Sequence 5, Application US/10124498  
; Publication No. US20030017983A1  
GENERAL INFORMATION:  
APPLICANT: Bertin, John  
APPLICANT: Wang, Weiye  
TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR  
FILE REFERENCE: 07334-367001  
CURRENT APPLICATION NUMBER: US/10/124,498  
CURRENT FILING DATE: 2002-04-17  
PRIOR APPLICATION NUMBER: 10/066,521  
PRIOR FILING DATE: 2002-01-31  
PRIOR APPLICATION NUMBER: 60/318,645  
PRIOR FILING DATE: 2001-09-10  
PRIOR APPLICATION NUMBER: 60/265,231  
PRIOR FILING DATE: 2001-01-31  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 4035  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(4032)  
US-10-124-498-5

Query Match 100.0%; Score 4035; DB 14; Length 4035;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGAGCAAAATCGCTCACTTTTCAGCTACGGGCTGCAATGGTGTCTATGAG 60  
DB 1 ATGAGAGAGCAAAATCGCTCACTTTTCAGCTACGGGCTGCAATGGTGTCTATGAG 60

QY 61 CTAGCAAGAGAAATTTTCAAGATTCAAGAAATTAACAAAGAAATTTTCAAGAAATG 120  
DB 61 CTAGCAAGAGAAATTTTCAAGATTCAAGAAATTAACAAAGAAATTTTCAAGAAATG 120

QY 121 ACCACATGCTCTATTCACAGATTGAAATCGAAATGCAAGTGAATGTCTGGCACTC 180  
DB 121 ACCACATGCTCTATTCACAGATTGAAATCGAAATGCAAGTGAATGTCTGGCACTC 180

QY 181 CTCTTGATGAGTATTTATGAGACATGCTGGCTGGGCTACGTCATTAGACATTTTGA 240  
DB 181 CTCTTGATGAGTATTTATGAGACATGCTGGCTGGGCTACGTCATTAGACATTTTGA 240

QY 241 AACATGAACCTGGCAACCTCTCGAGAAAGCACGGATGACATGAAAAATTTTCAAA 300  
DB 241 AACATGAACCTGGCAACCTCTCGAGAAAGCACGGATGACATGAAAAATTTTCAAA 300

QY 301 GCTATGAAACAAGAGTGGCCACAGCAGACAGAAACAAGAAATTTTCAAACT 360  
DB 301 GCTATGAAACAAGAGTGGCCACAGCAGACAGAAACAAGAAATTTTCAAACT 360

QY 361 ATGGAACAAGAGTGGCCACAGCAGACAGAAACAAGAAACAAGATGAGAGTGAACA 420  
DB 361 ATGGAACAAGAGTGGCCACAGCAGACAGAAACAAGAAACAAGATGAGAGTGAACA 420

QY 421 TGGGACTCAAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGTACGTCGCTAGT 480  
DB 421 TGGGACTCAAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGTACGTCGCTAGT 480

QY 481 TTTGAAAAACACTGCTGCTGACTGGCCGGAATGCAAGTGGCTGTGCTTTGATTC 540  
DB 481 TTTGAAAAACACTGCTGCTGACTGGCCGGAATGCAAGTGGCTGTGCTTTGATTC 540

QY 541 GACCGGTGGGCTTCGGCTTCGCAAGGTGTTCTGCAAGAAAGTCAAGAAATGGGAAA 600  
DB 541 GACCGGTGGGCTTCGGCTTCGCAAGGTGTTCTGCAAGAAAGTCAAGAAATGGGAAA 600

QY 601 TCGGCTAGCCAGAGAGATGCTGTGCTGGGCGCAAGGTGACCTTACAGGGAATG 660  
DB 601 TCGGCTAGCCAGAGAGATGCTGTGCTGGGCGCAAGGTGACCTTACAGGGAATG 660

QY 661 TTTCTCTACGTTCTTCTCCCGTTAGAGATGACGCGGAAGAGAGAGAGAGAGTGC 720  
DB 661 TTTCTCTACGTTCTTCTCCCGTTAGAGATGACGCGGAAGAGAGAGAGAGAGTGC 720

QY 721 ACAAGATTCAATCTCCAGGAGTGGCCAGACTCCAGGCTCCGTTGACGAGATCATGTC 780  
DB 721 ACAAGATTCAATCTCCAGGAGTGGCCAGACTCCAGGCTCCGTTGACGAGATCATGTC 780

QY 781 CGACCAAGAAAGGCTGTTTATCATGATGACGTTTCCATGACCTGGGCTCTGCTCTAAC 840  
DB 781 CGACCAAGAAAGGCTGTTTATCATGATGACGTTTCCATGACCTGGGCTCTGCTCTAAC 840

QY 841 AATGACAAAGAGCTTCGAAAGACTGGGCTGAGAGACAGCTCCGTTACACCTCAAGC 900  
DB 841 AATGACAAAGAGCTTCGAAAGACTGGGCTGAGAGACAGCTCCGTTACACCTCAAGC 900

QY 901 AGTCTGCTGAGAGAGTCTGCTCCCTGAGTCTTCTGATGCTACCGTCAAGAGAGT 960  
DB 901 AGTCTGCTGAGAGAGTCTGCTCCCTGAGTCTTCTGATGCTACCGTCAAGAGAGT 960

QY 961 GGCACAGAGAGCTCAAGTCAAGAGTCTGCTCCCGCTTACCTGTTAGTTAGAGAGATC 1020  
DB 961 GGCACAGAGAGCTCAAGTCAAGAGTCTGCTCCCGCTTACCTGTTAGTTAGAGAGATC 1020

QY 1021 TCCGGGAAACAAGATCACTTGTCTTGAAGCGCGGATTTGTAAGCATCAAGACA 1080

DB 1021 TCCGGGAAACAAGATCACTTGTCTTGAAGCGCGGATTTGTAAGCATCAAGACA 1080

QY 1081 CAAGGTTGGCTGGATCATGAAACAACCGTGAAGTCTCGAACAGTGCAGGTGCCGCC 1140  
DB 1081 CAAGGTTGGCTGGATCATGAAACAACCGTGAAGTCTCGAACAGTGCAGGTGCCGCC 1140

QY 1141 GTGGGCTCTCATCTGCTGAGCCCTGAGAGTGCAGAGACGTTGGGGAGAGGCTGCC 1200  
DB 1141 GTGGGCTCTCATCTGCTGAGCCCTGAGAGTGCAGAGACGTTGGGGAGAGGCTGCC 1200

QY 1201 CCTTTCAACAAAGCTCAAGGCTGACGCGCTTTGTTGTTATCATGACTCAACCT 1260  
DB 1201 CCTTTCAACAAAGCTCAAGGCTGACGCGCTTTGTTGTTATCATGACTCAACCT 1260

QY 1261 CGAGGCTGGTCCGGGCTGTCTCATCTGAGAGAAAGATTGCTGAAAGCGTTCTGC 1320  
DB 1261 CGAGGCTGGTCCGGGCTGTCTCATCTGAGAGAAAGATTGCTGAAAGCGTTCTGC 1320

QY 1321 CGTATGGCTGGAGAGTGGAAATGAGAGTCACTGTTGATGGTGAAGCATGATG 1380  
DB 1321 CGTATGGCTGGAGAGTGGAAATGAGAGTCACTGTTGATGGTGAAGCATGATG 1380

QY 1381 GTTCAAGAGACTCGGGAGTCTGAGCTCCGCTCTGTTTCAATGAACATCTTCCCA 1440  
DB 1381 GTTCAAGAGACTCGGGAGTCTGAGCTCCGCTCTGTTTCAATGAACATCTTCCCA 1440

QY 1441 GACAGCACTGTGAGAGTACTACCTTTTCAACCTCAAGTCTCCAGAGCTTGTGCC 1500  
DB 1441 GACAGCACTGTGAGAGTACTACCTTTTCAACCTCAAGTCTCCAGAGCTTGTGCC 1500

QY 1501 GCTTTGACTACGTTTGAAGGGGCTGGAATGAGGACGCTCTGCGCTGTAGT 1560  
DB 1501 GCTTTGACTACGTTTGAAGGGGCTGGAATGAGGACGCTCTGCGCTGTAGT 1560

QY 1561 GAGAAACAAGAGTCCATGAGCTTAAACAGGCAAGCTTCAATCACTCGCTTTGG 1620  
DB 1561 GAGAAACAAGAGTCCATGAGCTTAAACAGGCAAGCTTCAATCACTCGCTTTGG 1620

QY 1621 ATGAAGGTTCTTGTGCTGCTGCTGAGAGGAGAGTGAAGAGGCACTGAGAGTCTG 1680  
DB 1621 ATGAAGGTTCTTGTGCTGCTGCTGAGAGGAGAGTGAAGAGGCACTGAGAGTCTG 1680

QY 1681 CTGGGCTGTCCCGTTCGCGGGGTGAAGAGAGTTCGCACTGGGCTCTCTGTTG 1740  
DB 1681 CTGGGCTGTCCCGTTCGCGGGGTGAAGAGAGTTCGCACTGGGCTCTCTGTTG 1740

QY 1741 GGTGAGAGCTTAATGCAACCAACCCAGAGACACCTGAGCGCTTCACTGTCTTTTC 1800  
DB 1741 GGTGAGAGCTTAATGCAACCAACCCAGAGACACCTGAGCGCTTCACTGTCTTTTC 1800

QY 1801 GAGACTCAAGCAAAAGATTGTTGCTTGGCATTAACAGCTTCAAGAAATGTGGCTT 1860  
DB 1801 GAGACTCAAGCAAAAGATTGTTGCTTGGCATTAACAGCTTCAAGAAATGTGGCTT 1860

QY 1861 CCGATTAAACAAGACTGAGCTTGAATGACATCTTCTGCTCCGACCTGCGGAT 1920  
DB 1861 CCGATTAAACAAGACTGAGCTTGAATGACATCTTCTGCTCCGACCTGCGGAT 1920

QY 1921 TTGCGAAAAATTCGGGTGATGTCMAAGGATCTTCCCAAGATGAGTCCGCTGAGCA 1980  
DB 1921 TTGCGAAAAATTCGGGTGATGTCMAAGGATCTTCCCAAGATGAGTCCGCTGAGCA 1980

QY 1981 TGTCTGTGCTCCTCTATGATGTCGGGATTAAGACCTCTATGAGAGAGTGGGAAT 2040  
DB 1981 TGTCTGTGCTCCTCTATGATGTCGGGATTAAGACCTCTATGAGAGAGTGGGAAT 2040

QY 2041 TTTGCTCAATGCTTTGACCAACCAACCTGGGAGAGTGGGAGAGAGATC 2100  
DB 2041 TTTGCTCAATGCTTTGACCAACCAACCTGGGAGAGTGGGAGAGAGATC 2100

QY 2101 CTGACAGAGGGGCAATGAGAGCTGTGTGCAAGCTGAGGATCCCACTGCAAGATA 2160

Db 2101 CTGACGAGCGGGCCATGAAGACCTGTGTGCCAAGCTGAGGATCCCACTGCAAGATA 2160  
Qy 2161 CAGACCCCTGATGTTTAAAGAAATGACAGATTACCCCTGTGTGACAGACCTCTGAGAAATC 2220  
Db 2161 CAGACCCCTGATGTTTAAAGAAATGACAGATTACCCCTGTGTGACAGACCTCTGAGAAATC 2220  
Qy 2221 GTCATGGCCCAACCGTAACTTAAGATCCCTCACTTGGGAGGACCACTGAAGAGAG 2280  
Db 2221 GTCATGGCCCAACCGTAACTTAAGATCCCTCACTTGGGAGGACCACTGAAGAGAG 2280  
Qy 2281 GATGTAAAGATGGCGTGTGAAAGCCTTAAACAACCAAAATGTTTGTGAGTCTTTGAGG 2340  
Db 2281 GATGTAAAGATGGCGTGTGAAAGCCTTAAACAACCAAAATGTTTGTGAGTCTTTGAGG 2340  
Qy 2341 CTGGAATGTGTGATGATGACCAATGCTGTACCTGGAAGATCTCCCAATCTTTGAGC 2400  
Db 2341 CTGGAATGTGTGATGATGACCAATGCTGTACCTGGAAGATCTCCCAATCTTTGAGC 2400  
Qy 2401 TCCCCAGCGCTGAATCTCTGAGCGTGGCAGGAAACAAGGTGACAGACCGAGGATATG 2460  
Db 2401 TCCCCAGCGCTGAATCTCTGAGCGTGGCAGGAAACAAGGTGACAGACCGAGGATATG 2460  
Qy 2461 CCTCTCAGTATGCTCTTGAAGATCTCCAGTGCCTGACAGAGCTGATCTGAGAGAC 2520  
Db 2461 CCTCTCAGTATGCTCTTGAAGATCTCCAGTGCCTGACAGAGCTGATCTGAGAGAC 2520  
Qy 2521 TGTGCAATCAACGCCACCGGCTTCCAGAGTCTGAGCTCAGCCCTGTGACAGACCGAGC 2580  
Db 2521 TGTGCAATCAACGCCACCGGCTTCCAGAGTCTGAGCTCAGCCCTGTGACAGACCGAGC 2580  
Qy 2581 TTGACACACCTGTGCTATCCAAACAAGCCTGGGGAACGAGGTGAATCTACCTGCT 2640  
Db 2581 TTGACACACCTGTGCTATCCAAACAAGCCTGGGGAACGAGGTGAATCTACCTGCT 2640  
Qy 2641 CGATCCATGAGGCTTCCCACTGATGTGACAGAGCTGATCTGAATCAAGTCCACCTG 2700  
Db 2641 CGATCCATGAGGCTTCCCACTGATGTGACAGAGCTGATCTGAATCAAGTCCACCTG 2700  
Qy 2701 GACACGCGCTGCTGTGTTTCTTGCACCTTGGCTTAATGAGTCACTGAGCTGACGAC 2760  
Db 2701 GACACGCGCTGCTGTGTTTCTTGCACCTTGGCTTAATGAGTCACTGAGCTGACGAC 2760  
Qy 2761 CTGACCGCTTAAGATTAACCCCTGTGAAAGCAATGCGGTGAAGCTTGTGCGAGGTCA 2820  
Db 2761 CTGACCGCTTAAGATTAACCCCTGTGAAAGCAATGCGGTGAAGCTTGTGCGAGGTCA 2820  
Qy 2821 AGAGAACCACTTGTCTATCTCAGAGACCTGAGTGTGTAAGTGTCTATCTACCCGCG 2880  
Db 2821 AGAGAACCACTTGTCTATCTCAGAGACCTGAGTGTGTAAGTGTCTATCTACCCGCG 2880  
Qy 2881 TGTCTGTGAGAGTCTGTCTGTGTGATCTTGCAGAGACAGACCTGAAAGCTGTGATCTC 2940  
Db 2881 TGTCTGTGAGAGTCTGTCTGTGTGATCTTGCAGAGACAGACCTGAAAGCTGTGATCTC 2940  
Qy 2941 ACGGACAAATGCGCTGTGAGAGCGTGTGAGTGTGCTGTGAGAGGAGCTGAAGCAAAAG 3000  
Db 2941 ACGGACAAATGCGCTGTGAGAGCGTGTGAGTGTGCTGTGAGAGGAGCTGAAGCAAAAG 3000  
Qy 3001 AACAGTGTCTGACAGAGACTCGGCTTGAAGCATGTGACCTGATCTGATGCTGTGAG 3060  
Db 3001 AACAGTGTCTGACAGAGACTCGGCTTGAAGCATGTGACCTGATCTGATGCTGTGAG 3060  
Qy 3061 GCACCTCTCTTGGCCCTTCTGCAACCGGCACTGACAGCTTAACCTGTGTGAGAT 3120  
Db 3061 GCACCTCTCTTGGCCCTTCTGCAACCGGCACTGACAGCTTAACCTGTGTGAGAT 3120  
Qy 3121 AACTTCAGTCCCAAGAGATGATGAAGCTGTGTGAGCTTGGCTTGTGCTGCTGCTTAC 3180  
Db 3121 AACTTCAGTCCCAAGAGATGATGAAGCTGTGTGAGCTTGGCTTGTGCTGCTGCTTAC 3180  
Qy 3181 TTACAGATTAATGGGCTGTGAAATGAGCATGCTGCTGCAATTAAGAGAGCTGCTGAG 3240  
Db 3181 TTACAGATTAATGGGCTGTGAAATGAGCATGCTGCTGCAATTAAGAGAGCTGCTGAG 3240

Qy 3241 GAAGTGCAGTACTCAAGCCCCGAGTGTGATTAAGACGCTAGTGTGATCTTTGATGA 3300  
Db 3241 GAAGTGCAGTACTCAAGCCCCGAGTGTGATTAAGACGCTAGTGTGATCTTTGATGA 3300  
Qy 3301 GATGACCGACCAAAATAGACTTAATCTTCCGCTCCCTGAAAGCGGAGTGCATGT 3360  
Db 3301 GATGACCGACCAAAATAGACTTAATCTTCCGCTCCCTGAAAGCGGAGTGCATGT 3360  
Qy 3361 GCTTGTGTGTGGGATGAACCCAGAGCAAAAGCTGTGTGCTTGTGCTGAGAC 3420  
Db 3361 GCTTGTGTGTGGGATGAACCCAGAGCAAAAGCTGTGTGCTTGTGCTGAGAC 3420  
Qy 3421 TTCAAGACAGTACCAAGTTTGGCAAGTCTCTGCTGCGCAACGCAAAATGTGAGTCC 3480  
Db 3421 TTCAAGACAGTACCAAGTTTGGCAAGTCTCTGCTGCGCAACGCAAAATGTGAGTCC 3480  
Qy 3481 CAGAGAGTTGACCAAGTGTGACAGAGCTCCCGCAACCATGAGCAGCAGCAACAA 3540  
Db 3481 CAGAGAGTTGACCAAGTGTGACAGAGCTCCCGCAACCATGAGCAGCAGCAACAA 3540  
Qy 3541 CAAGTAAATGTTGATGTTGATTAATCCGAGCTGTGTGAAATGCTGAGCTGAA 3600  
Db 3541 CAAGTAAATGTTGATGTTGATTAATCCGAGCTGTGTGAAATGCTGAGCTGAA 3600  
Qy 3601 GGGCTTGTATCCAAACAGTGTGATTAATCCAGAGCTGTGTGAAATGCTGAGCTGAA 3660  
Db 3601 GGGCTTGTATCCAAACAGTGTGATTAATCCAGAGCTGTGTGAAATGCTGAGCTGAA 3660  
Qy 3661 GAGCTGAGCTGAGGAGCTGTGTGCAACAGTGTGATGACCAACAGCGTGTGTCTG 3720  
Db 3661 GAGCTGAGCTGAGGAGCTGTGTGCAACAGTGTGATGACCAACAGCGTGTGTCTG 3720  
Qy 3721 CACTGGAACCGCTGTGAGCTTGAAGGCTGTGTCTTAAACAGTGTGATGACCAACAG 3780  
Db 3721 CACTGGAACCGCTGTGAGCTTGAAGGCTGTGTCTTAAACAGTGTGATGACCAACAG 3780  
Qy 3781 GTGTCTGTGTCACTGTGAGAGCGGCTGTGAGAGGCTGTGTCTTAAACAGTGTGATG 3840  
Db 3781 GTGTCTGTGTCACTGTGAGAGCGGCTGTGAGAGGCTGTGTCTTAAACAGTGTGATG 3840  
Qy 3841 CACAGCGTGTGAGCTGTGTGCACTGTGAGAGCGGCTGTGAGAGGCTGTGTCTTAA 3900  
Db 3841 CACAGCGTGTGAGCTGTGTGCACTGTGAGAGCGGCTGTGAGAGGCTGTGTCTTAA 3900  
Qy 3901 GCTGATGACCAACAGCGTGTGTCTGTGATGAGAGCGGCTGTGAGAGGCTGTGTG 3960  
Db 3901 GCTGATGACCAACAGCGTGTGTCTGTGATGAGAGCGGCTGTGAGAGGCTGTGTG 3960  
Qy 3961 TCACACAGTGTGATGACCAACAGCGGCTGTGTGATGAGAGCGGCTGTGAGAG 4020  
Db 3961 TCACACAGTGTGATGACCAACAGCGGCTGTGTGATGAGAGCGGCTGTGAGAG 4020  
Qy 4021 GGGCTGTGTCTTAA 4035  
Db 4021 GGGCTGTGTCTTAA 4035

RESULT 2  
US-10-066-521-5  
Sequence 5, Application US/10066521  
Publication No. US20030027757A1  
GENERAL INFORMATION:  
APPLICANT: Bertin, John  
APPLICANT: Wang, Weiye  
APPLICANT: Blacher, Maria  
TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR  
FILE REFERENCE: 07334-334001  
CURRENT APPLICATION NUMBER: US/10/066, 521  
CURRENT FILING DATE: 2002-06-25  
PRIOR APPLICATION NUMBER: 60/318, 645  
PRIOR FILING DATE: 2001-09-10

;; PRIOR APPLICATION NUMBER: 60/265,231  
;; PRIOR FILING DATE: 2001-01-31  
;; NUMBER OF SEQ ID NOS: 25  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 5  
;; LENGTH: 4035  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (1)...(4032)  
US-10-066-521-5

Query Match 100.0%; Score 4035; DB 14; Length 4035;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAGAGCAAAATGCTCACTTTTCAGCTACGGGCTGCAATGCTGCTCTATGAG 60  
DB 1 ATGGAAGAGCAAAATGCTCACTTTTCAGCTACGGGCTGCAATGCTGCTCTATGAG 60  
QY 61 CTAGCAAGAGAAATTTCAAGCATTTCAAGAAATTAACAAAGAAATTTCAAGAAATCG 120  
DB 61 CTAGCAAGAGAAATTTCAAGCATTTCAAGAAATTAACAAAGAAATTTCAAGAAATCG 120  
QY 121 ACCACATGCTCTATTCACACAGTTTGAATTCAGAAATGCAAGTGAATGCTGACATC 180  
DB 121 ACCACATGCTCTATTCACACAGTTTGAATTCAGAAATGCAAGTGAATGCTGACATC 180  
QY 181 CTCTTGCAATGATATTAATGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
DB 181 CTCTTGCAATGATATTAATGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
QY 241 AACATGAACCTGCGAACCCTCTCGGAGAAAGCAAGGATGACATGAAAAATTTCAAA 300  
DB 241 AACATGAACCTGCGAACCCTCTCGGAGAAAGCAAGGATGACATGAAAAATTTCAAA 300  
QY 301 GCTATGAAACAAGAGTGCACACAGCAGACAGACAGAAATTTCAAGAAATTTCAAG 360  
DB 301 GCTATGAAACAAGAGTGCACACAGCAGACAGACAGAAATTTCAAGAAATTTCAAG 360  
QY 361 ATGGAACAAGAGTGCACACAGCAGACAGACAGAAATTTCAAGAAATTTCAAG 420  
DB 361 ATGGAACAAGAGTGCACACAGCAGACAGACAGAAATTTCAAGAAATTTCAAG 420  
QY 421 TGGGCTCAAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGATGACGCTGATAGT 480  
DB 421 TGGGCTCAAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGATGACGCTGATAGT 480  
QY 481 TTTGAAAAACATGCTGCTGACCTGCGGAAATGCAAAAGTTGCTGCTGCTGCTGCTG 540  
DB 481 TTTGAAAAACATGCTGCTGACCTGCGGAAATGCAAAAGTTGCTGCTGCTGCTGCTG 540  
QY 541 GACCGGTGGGCTTCGCGCTGCGACAGGCTGCTGCAAGAAATGCAAGAAATTTG 600  
DB 541 GACCGGTGGGCTTCGCGCTGCGACAGGCTGCTGCAAGAAATGCAAGAAATTTG 600  
QY 601 TCGGCTCAAGCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
DB 601 TCGGCTCAAGCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
QY 661 TTTCTCTACGCTTTCTTCTCCCGTTAGAGAGATGACAGCGAGAGAGAGAGAGAGT 720  
DB 661 TTTCTCTACGCTTTCTTCTCCCGTTAGAGAGATGACAGCGAGAGAGAGAGAGAGT 720  
QY 721 ACAGAGTTCATCTCAGGAGAGTGCACAGCTCCAGGCTCCGCTGACGAGATCATGCTC 780  
DB 721 ACAGAGTTCATCTCAGGAGAGTGCACAGCTCCAGGCTCCGCTGACGAGATCATGCTC 780  
QY 781 CGACCAAGAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
DB 781 CGACCAAGAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840

QY 841 AATGACAAAGCTCTGCAAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
DB 841 AATGACAAAGAGCTCTGCAAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
QY 901 AGTCTGCTGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
DB 901 AGTCTGCTGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
QY 961 GGCACAGAGAGAGCTCAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1020  
DB 961 GGCACAGAGAGAGCTCAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1020  
QY 1021 TCCGAGAGAAAGATTCACCTTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
DB 1021 TCCGAGAGAAAGATTCACCTTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
QY 1081 CAAGGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
DB 1081 CAAGGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
QY 1141 GTGGGCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
DB 1141 GTGGGCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
QY 1201 CCTTTCACCAAGAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
DB 1201 CCTTTCACCAAGAGAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
QY 1261 CGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
DB 1261 CGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
QY 1321 CGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
DB 1321 CGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
QY 1381 GTTCAAGAGAGCTGAG 1440  
DB 1381 GTTCAAGAGAGCTGAG 1440  
QY 1441 GACAGCACTGCTGAG 1500  
DB 1441 GACAGCACTGCTGAG 1500  
QY 1501 GCTTGTACTACAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
DB 1501 GCTTGTACTACAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
QY 1561 GAGAGACAAAGAGAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
DB 1561 GAGAGACAAAGAGAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
QY 1621 ATGAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680  
DB 1621 ATGAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680  
QY 1681 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740  
DB 1681 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740  
QY 1741 GGTGAGAGAGCTTAATGACCAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800  
DB 1741 GGTGAGAGAGCTTAATGACCAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800  
QY 1801 GAGACTCAAGCAAAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
DB 1801 GAGACTCAAGCAAAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
QY 1861 CCGATTAACCAAGAT 1920  
DB 1861 CCGATTAACCAAGAT 1920  
QY 1921 TTGCGAAAAATTCGGGTGAGATGTCAAAGAGAGATCTTCCAAAGAGATGAGTCCGCTGAGGCA 1980

Db 1921 TTGGGAAAAATTGGGGTGAATGTCAAAGGATCTTCCAAAGATGATGCCGTGAGGCA 1980  
Qy 1981 TGTCTGTGTGTCTCTATGATGTGGGATTAAGACCTTATTGAGGAGAGTGGAGAT 2040  
Db 1981 TGTCTGTGTGTCTCTATGATGTGGGATTAAGACCTTATTGAGGAGAGTGGAGAT 2040  
Qy 2041 TGTCTGTGTGTCTCTATGATGTGGGATTAAGACCTTATTGAGGAGAGTGGAGAT 2100  
Db 2041 TGTCTGTGTGTCTCTATGATGTGGGATTAAGACCTTATTGAGGAGAGTGGAGAT 2100  
Qy 2101 CTGACAGAGCGGGCCATGAAAGACCTGTGTGCAAGCTGAGGATCCCACTGCAAGATA 2160  
Db 2101 CTGACAGAGCGGGCCATGAAAGACCTGTGTGCAAGCTGAGGATCCCACTGCAAGATA 2160  
Qy 2161 CAGACCCCTGATGTTTGAAGATGACAGATTAACCTGTGTGAGACCTTGTGAGATC 2220  
Db 2161 CAGACCCCTGATGTTTGAAGATGACAGATTAACCTGTGTGAGACCTTGTGAGATC 2220  
Qy 2221 GTGATGGCCAAACGTTAAGATCCCTCACTTGGGAGGACCACTGAAAGAGAG 2280  
Db 2221 GTGATGGCCAAACGTTAAGATCCCTCACTTGGGAGGACCACTGAAAGAGAG 2280  
Qy 2281 GATGTAAAGATGCGGTGTGAAGCCTTAAACACCCAAATGTTGTGAGTCTTGAAG 2340  
Db 2281 GATGTAAAGATGCGGTGTGAAGCCTTAAACACCCAAATGTTGTGAGTCTTGAAG 2340  
Qy 2341 CTGATGTGTGTGATGACCAATGCTGTGTAAGATCTCCAAATCTTGAAGAC 2400  
Db 2341 CTGATGTGTGTGATGACCAATGCTGTGTAAGATCTCCAAATCTTGAAGAC 2400  
Qy 2401 TCCCCAGCTGAATCTGTGAGCCTGTGAGGAAACAGGTGACAGACCAAGGATTAATG 2460  
Db 2401 TCCCCAGCTGAATCTGTGAGCCTGTGAGGAAACAGGTGACAGACCAAGGATTAATG 2460  
Qy 2461 CCTCTCAGTGATGCTGTGAGAGTCTCCAGTGTGCTGCAAGCTGATATCTGAGAG 2520  
Db 2461 CCTCTCAGTGATGCTGTGAGAGTCTCCAGTGTGCTGCAAGCTGATATCTGAGAG 2520  
Qy 2521 TGTGATCAACAGCAAGCGGTTGCCAGAGTGTGCTGAGCCTGTGCAAGCGGAGC 2580  
Db 2521 TGTGATCAACAGCAAGCGGTTGCCAGAGTGTGCTGAGCCTGTGCAAGCGGAGC 2580  
Qy 2581 TTGACACACCTGTGCTATCCACACAGCCTGGGAAACAGAGTGTAAATCTACTGTGT 2640  
Db 2581 TTGACACACCTGTGCTATCCACACAGCCTGGGAAACAGAGTGTAAATCTACTGTGT 2640  
Qy 2641 CGATCATGAGGCTTCCCACTGTAGTCTGACAGAGGCTGATGCTGAATCAGTCCAGC 2700  
Db 2641 CGATCATGAGGCTTCCCACTGTAGTCTGACAGAGGCTGATGCTGAATCAGTCCAGC 2700  
Qy 2701 GACACGAGTGTGTGTCTGTGCACTTGTGCTTAAAGTGAATCTCAATGCTGACGAC 2760  
Db 2701 GACACGAGTGTGTGTCTGTGCACTTGTGCTTAAAGTGAATCTCAATGCTGACGAC 2760  
Qy 2761 CTGAGCCTTAAAGTGAATGACCTGTGAAAGCAATGCTGAAAGCTTGTGCAAGCTCATG 2820  
Db 2761 CTGAGCCTTAAAGTGAATGACCTGTGAAAGCAATGCTGAAAGCTTGTGCAAGCTCATG 2820  
Qy 2821 AGAGAACCATCTTGTGATCTCCAGAGCCTGAGGTTGTAAGTGTCAATCTCAAGCGGCG 2880  
Db 2821 AGAGAACCATCTTGTGATCTCCAGAGCCTGAGGTTGTAAGTGTCAATCTCAAGCGGCG 2880  
Qy 2881 TGTGTGAGAGTGTGTCTGTGTGATCTGAGAGACAGACCTGAAAGAGCTGTGATCTC 2940  
Db 2881 TGTGTGAGAGTGTGTCTGTGTGATCTGAGAGACAGACCTGAAAGAGCTGTGATCTC 2940  
Qy 2941 ACAGCAATGCTGTGTGTGAGAGGCTGTGCTGTGCTGTGAGAGGAGCTGAAGAAAG 3000  
Db 2941 ACAGCAATGCTGTGTGTGAGAGGCTGTGCTGTGCTGTGAGAGGAGCTGAAGAAAG 3000  
Qy 3001 AACAGTGTGTGACAGAGCTGGGTTGAAGGCAATGAGCACTTGTGATGTGTGAG 3060  
Db 3001 AACAGTGTGTGACAGAGCTGGGTTGAAGGCAATGAGCACTTGTGATGTGTGAG 3060

Db 3001 AACAGTGTGTGACAGAGCTGGGTTGAAGGCAATGAGCACTTGTGATGTGTGAG 3060  
Qy 3061 GCATCTCTCTTGGCCCTTTCCTGCAACGGGATCTGACAGCTTAAACCTGTGTGAGAT 3120  
Db 3061 GCATCTCTCTTGGCCCTTTCCTGCAACGGGATCTGACAGCTTAAACCTGTGTGAGAT 3120  
Qy 3121 AACTTCACTCCAAAGATGAGAGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 3180  
Db 3121 AACTTCACTCCAAAGATGAGAGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 3180  
Qy 3181 TTACAGATTAATGTGCTGTGAAATGCAAGTACCTGTGTGCAATTAAGAGCTGTGAG 3240  
Db 3181 TTACAGATTAATGTGCTGTGAAATGCAAGTACCTGTGTGCAATTAAGAGCTGTGAG 3240  
Qy 3241 GAGTGTGAGTGTGTGAGAGGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTG 3300  
Db 3241 GAGTGTGAGTGTGTGAGAGGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTG 3300  
Qy 3301 GATGACCGACCAAAATGAGACTTACTTCCGGCTCCCTGAAAGCGGGCATGTGCTGTG 3360  
Db 3301 GATGACCGACCAAAATGAGACTTACTTCCGGCTCCCTGAAAGCGGGCATGTGCTGTG 3360  
Qy 3361 GCTTGTGTGTGTGTGAGTGAACCCAGAGCAAGAGCTGTGTGCTGTGTGCTGTGTG 3420  
Db 3361 GCTTGTGTGTGTGTGAGTGAACCCAGAGCAAGAGCTGTGTGCTGTGTGCTGTGTG 3420  
Qy 3421 TTCAAGAGAGTGAACGATTTGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 3480  
Db 3421 TTCAAGAGAGTGAACGATTTGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 3480  
Qy 3481 CAGAGAGTTGAACAAGTGAAGAGGCTCCCGGACCCAGTGTGAGAGGAGAGAGAGAG 3540  
Db 3481 CAGAGAGTTGAACAAGTGAAGAGGCTCCCGGACCCAGTGTGAGAGGAGAGAGAGAG 3540  
Qy 3541 CAAGATTAATGTTGATGTGTGATTAATCCGAGCTGTGTGCTGTGCTGTGCTGTGCT 3600  
Db 3541 CAAGATTAATGTTGATGTGTGATTAATCCGAGCTGTGTGCTGTGCTGTGCTGTGCT 3600  
Qy 3601 GGGCTTGTGATCAACAGTGTGATGATGACCAAGAGGATGAGCTGTGCTGTGCTGTG 3660  
Db 3601 GGGCTTGTGATCAACAGTGTGATGATGACCAAGAGGATGAGCTGTGCTGTGCTGTG 3660  
Qy 3661 GAGCTGAGCTGAGGAGGCTGTGTGCTGCAAGTGTGATGACCAAGAGGATGAGCTGT 3720  
Db 3661 GAGCTGAGCTGAGGAGGCTGTGTGCTGCAAGTGTGATGACCAAGAGGATGAGCTGT 3720  
Qy 3721 CACTGTGAGAGGCTGTGAGGCTGTGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 3780  
Db 3721 CACTGTGAGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGCTGTGCTGTGCTGTG 3780  
Qy 3781 GTGTCTGTGTGATGAGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 3840  
Db 3781 GTGTCTGTGTGATGAGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 3840  
Qy 3841 CACAGCGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 3900  
Db 3841 CACAGCGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 3900  
Qy 3901 GGTGATGACCAAGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 3960  
Db 3901 GGTGATGACCAAGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 3960  
Qy 3961 TCCAAAGTGTGATGACCAAGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 4020  
Db 3961 TCCAAAGTGTGATGACCAAGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 4020  
Qy 4021 GGGCTGTGTGTGCTTAA 4035  
Db 4021 GGGCTGTGTGTGCTTAA 4035

RESULT 3  
US-10-860-761-3



Sequence 3, Application US/10860761  
; Publication No. US20040248755A1  
; GENERAL INFORMATION:  
; APPLICANT: WYETH  
; TITLE OF INVENTION: METHODS FOR SCREENING INHIBITORS OF APOPTOSIS  
; FILE REFERENCE: AM101318  
; CURRENT APPLICATION NUMBER: US/10/860,761  
; CURRENT FILING DATE: 2004-06-03  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO: 3  
; LENGTH: 3885  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(3603)  
US-10-860-761-3

Query Match 78.5%; Score 3166.2; DB 20; Length 3885;  
Best Local Similarity 96.2%; Pred. No. 0;  
Matches 3304; Conservative 0; Mismatches 8; Indels 123; Gaps 1;

```
QY 1 ATGGAAGAGCAAAATCGCTCACTTTTCCAGCTACGCGCTGCAATGGTGTCTATGAG 60
DB 154 ATGGAAGAGCAAAATCGCTCACTTTTCCAGCTACGCGCTGCAATGGTGTCTATGAG 213
QY 61 CTAGCAAGGAAGATTTCAGCATTCAGAAATTACTTAAGAAAGAAATCTTCAGAAATCG 120
DB 214 CTAGCAAGGAAGATTTCAGCATTCAGAAATTACTTAAGAAAGAAATCTTCAGAAATCG 273
QY 121 ACCACATGCTCTATTCACAGTTTGAATTCAGAAATGCAAGTGAATGTCTGCACTC 180
DB 274 ACCACATGCTCTATTCACAGTTTGAATTCAGAAATGCAAGTGAATGTCTGCACTC 333
QY 181 CTCTTGCAATGATATTATGAGCATCGCTGCGCTGCGCTACGTCATTAGCATCTTTGAA 240
DB 334 CTCTTGCAATGATATTATGAGCATCGCTGCGCTGCGCTACGTCATTAGCATCTTTGAA 393
QY 241 AACATGAACCTGCGCAACCTCTCGGAGAAAGGACGGGATGACATGAAA----- 289
DB 394 AACATGAACCTGCGCAACCTCTCGGAGAAAGGACGGGATGACATGAAAAGCATTCACCA 453
QY 290 ----- 289
DB 454 GAAGATCTGAAGCAAGATGATGACCAAGGACCAAGCAAGAAAGTGCAGAAAT 513
QY 290 -----AAATTTCA 297
DB 514 TCACAAAGCTGTGCAACAAGATAGTGCACAGCTGCAGACAAAAGAAACAGAAATTTCA 573
QY 298 CAAGCTATGAAACAAGAGGTGCGCACAGCAGACAGACAGAAACAAGAAATTTCAAA 357
DB 574 CAAGCTATGAAACAAGAGGTGCGCACAGCAGACAGACAGAAACAAGAAATTTCAAA 633
QY 358 GCTATGAAACAAGAGGTGCGCACAGCAGACAGACAGAAACAAGACATGAGAGTAC 417
DB 634 GCTATGAAACAAGAGGTGCGCACAGCAGACAGACAGAAACAAGACATGAGAGTAC 693
QY 418 ACATGGGACTACAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGATGTACGTCT 477
DB 694 ACATGGGACTACAGAGTCAAGTATGACAAATTCGCTGAGAGAGAGATGTACGTCT 753
QY 478 AGTTTGAACCACTGCTGTGACCTGCGCGGAAATGCAAGTGGCTGGTCTTTTAT 537
DB 754 AGTTTGAACCACTGCTGTGACCTGCGCGGAAATGCAAGTGGCTGGTCTTTTAT 813
QY 538 TCAGACCGGTGGGCTTCCGCGCTGCGACGCGTGTCTGCAAGGAAATGCAAGATTTGG 597
DB 814 TCAGACCGGTGGGCTTCCGCGCTGCGACGCGTGTCTGCAAGGAAATGCAAGATTTGG 873
QY 598 AAATCGGCTTACCCAGCAAGATCGTGTGTGCTGGCGCAAGGTGACCTTACCAAGGA 657
DB 657 AAATCGGCTTACCCAGCAAGATCGTGTGTGCTGGCGCAAGGTGACCTTACCAAGGA 657
```

```
DB 874 AAATCGGCTTACCCAGCAAGATCGTGTGTGCTGGCGCAAGGTGACCTTACCAAGGA 933
QY 658 ATGTTCTCTACAGCTCTTCTCTCCCGTTAGAGAGATGACGCGAAGAAAGAGAGAGT 717
DB 934 ATGTTCTCTACAGCTCTTCTCTCCCGTTAGAGAGATGACGCGAAGAAAGAGAGAGT 993
QY 718 GTCCAGAGTTCAATCTCCAGGAGTGGCCAGACTCCAGGCTCCGTTGACGAGATCATG 777
DB 994 GTCCAGAGTTCAATCTCCAGGAGTGGCCAGACTCCAGGCTCCGTTGACGAGATCATG 1053
QY 778 TCCCGACCAAGAAAGCTGTGTTTATCATATGACGCTTTGATGACCTGGCTCTGCTC 837
DB 1054 TCCCGACCAAGAAAGCTGTGTTTATCATATGACGCTTTGATGACCTGGCTCTGCTC 1113
QY 838 AACATGACACAAAGCTCTGCAAGATGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAG 897
DB 1114 AACATGACACAAAGCTCTGCAAGATGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1173
QY 898 CGCAGTCTGTGAGAAAGTCTGCTCTGAGTCTTCTGATGTCACCGTCAAGAGAC 957
DB 1174 CGCAGTCTGTGAGAAAGTCTGCTCTGAGTCTTCTGATGTCACCGTCAAGAGAC 1233
QY 958 GTGGGCAACAGAAAGCTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTTAGTAAAGGA 1017
DB 1234 GTGGGCAACAGAAAGCTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTTAGTAAAGGA 1293
QY 1018 ATCTCCGGGGAAACAAAGATTCATCTGCTCCTTGAAGCGGAGATTTGAGAGATCAGAA 1077
DB 1294 ATCTCCGGGGAAACAAAGATTCATCTGCTCCTTGAAGCGGAGATTTGAGAGATCAGAA 1353
QY 1078 ACACAAAGGTTGCGTGCATCATGAAACACCGTGAAGTGTCTGACACAGTGCAGTGC 1137
DB 1354 ACACAAAGGTTGCGTGCATCATGAAACACCGTGAAGTGTCTGACACAGTGCAGTGC 1413
QY 1138 GCCGTGGGCTCTCTCATCTGCGTGGCCCTGACAGTGAAGAGACGTGTGGGGAGAGGCTC 1197
DB 1414 GCCGTGGGCTCTCTCATCTGCGTGGCCCTGACAGTGAAGAGACGTGTGGGGAGAGGCTC 1473
QY 1198 GCCCCTTCAACAAAGCTCAACAGGCTGCAACGCGCTTTTGTGTTCAATCAAGTCAAC 1257
DB 1474 GCCCCTTCAACAAAGCTCAACAGGCTGCAACGCGCTTTTGTGTTCAATCAAGTCAAC 1533
QY 1258 CTTGAGAGCGTGTGTCGCGCGCTGTCTCAATCTGAGAGAAAGATGTCTGAAAGCGCTTC 1317
DB 1534 CTTGAGAGCGTGTGTCGCGCGCTGTCTCAATCTGAGAGAAAGATGTCTGAAAGCGCTTC 1593
QY 1318 TGCCGTATGCTGTGAGAGAGTGTGAATAGAAAGTCAAGTGTGATGAGTGAACCTC 1377
DB 1594 TGCCGTATGCTGTGAGAGAGTGTGAATAGAAAGTCAAGTGTGATGAGTGAACCTC 1653
QY 1378 ATGATTCAAGGACTCGGAGAGTCTGAGCTCGTCTGTTTCAATGAACATCTTCTC 1437
DB 1554 ATGATTCAAGGACTCGGAGAGTCTGAGCTCGTCTGTTTCAATGAACATCTTCTC 1713
QY 1438 CCAGACAGCCACTGTGAGAGATCTACACCTTCTTCAACCTCACTGTCACAGACTTGT 1497
DB 1714 CCAGACAGCCACTGTGAGAGATCTACACCTTCTTCAACCTCACTGTCACAGACTTGT 1773
QY 1498 GCCGCTTGTATCTACGTTTGAAGAGGCTTGAATGAGACAGTCTCTGCTCTGTAC 1557
DB 1774 GCCGCTTGTATCTACGTTTGAAGAGGCTTGAATGAGACAGTCTCTGCTCTGTAC 1833
QY 1558 GTTGAAGAAACAAAGAGTGCATGAGACTTAAACAGGACAGCTTCAATATCACTCGCT 1617
DB 1834 GTTGAAGAAACAAAGAGTGCATGAGACTTAAACAGGACAGCTTCAATATCACTCGCT 1893
QY 1618 TGAATGAAGGCTTTCTTGTGTGCTGTGAGCGAAGACGTAAAGAGCCACTGAGAGT 1677
DB 1894 TGAATGAAGGCTTTCTTGTGTGCTGTGAGCGAAGACGTAAAGAGCCACTGAGAGT 1953
QY 1678 CTGCTGGGCTGTCCCGTTTCCCTGGGGGTGAAGCAAGCTTCTGCACTGGGTCTCTCTG 1737
DB 1954 CTGCTGGGCTGTCCCGTTTCCCTGGGGGTGAAGCAAGCTTCTGCACTGGGTCTCTCTG 2013
```

1738 TTGGGTGACAGACCTTAATGCGACACCCAGAGAGACACCTGAGAGCCTTCCACTGTCTT 1797  
1739 TTGGGTGACAGACCTTAATGCGACACCCAGAGAGACACCTGAGAGCCTTCCACTGTCTT 2073  
1798 TTGAGACTCAAGACCAAGAGTTTGTTCGCTTGGCATTAACAGCTTCCAGAGAGTGTG 1857  
2074 TTGAGACTCAAGACCAAGAGTTTGTTCGCTTGGCATTAACAGCTTCCAGAGAGTGTG 2133  
1858 CTTCCGATTAAACAGAACTTGAATGATCTTCTTCTGCTCCAGACCTGTCCG 1917  
2134 CTTCCGATTAAACAGAACTTGAATGATCTTCTTCTGCTCCAGACCTGTCCG 2193  
1918 TATTTGCGGAAATTCGCGGTGATGTCAGAAAGGATTTCCCAAGATAGTCGCTGAG 2253  
2194 TATTTGCGGAAATTCGCGGTGATGTCAGAAAGGATTTCCCAAGATAGTCGCTGAG 2253  
1978 GCATGCTCTGTGCTCCCTCTATGATGCGGATTAAGACCTTCATTAAGAGAGTGGAA 2037  
2254 GCATGCTCTGTGCTCCCTCTATGATGCGGATTAAGACCTTCATTAAGAGAGTGGAA 2313  
2038 GATTTCTGCTCATGCTTGGCAACCCACACCTGCGGACCTGAGACCTGGGACAGAC 2097  
2314 GATTTCTGCTCATGCTTGGCAACCCACACCTGCGGACCTGAGACCTGGGACAGAC 2373  
2098 ATCTGACAGAGCGGCGCATGAAAGACCTGTGTGCCAAGCTGAGGATCCACCTGACAG 2157  
2374 ATCTGACAGAGCGGCGCATGAAAGACCTGTGTGCCAAGCTGAGGATCCACCTGACAG 2433  
2158 ATACAGACCTGATGTTTGAATGACAGATTAACCTGCTGTGTCAGACCTCTGGAGA 2217  
2434 ATACAGACCTGATGTTTGAATGACAGATTAACCTGCTGTGTCAGACCTCTGGAGA 2493  
2218 ATGCTATGCGCAACCGTAACCTTAAGATCCCTCAACTTGGAGAGCACCCACTGAGAA 2277  
2494 ATGCTATGCGCAACCGTAACCTTAAGATCCCTCAACTTGGAGAGCACCCACTGAGAA 2553  
2278 GAGGATGTAAGATGCGGTGTGAAGCCTTAAACACCCAAATGTTTGTGAGTCTTGG 2337  
2554 GAGGATGTAAGATGCGGTGTGAAGCCTTAAACACCCAAATGTTTGTGAGTCTTGG 2613  
2338 AGGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2397  
2614 AGGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2673  
2398 ACCCTCCCAAGCCTGAATCTCTGAGCCTGCGAGGAAACAGAGTGAAGACAGAGGAGTA 2457  
2674 ACCCTCCCAAGCCTGAATCTCTGAGCCTGCGAGGAAACAGAGTGAAGACAGAGGAGTA 2733  
2458 ATGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2517  
2734 AGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2793  
2518 GACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2577  
2794 GACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2853  
2578 AGCTTGACACACCTGCTTATCAACACAGCCTGCGGAGAGAGAGTGAATCTACTG 2637  
2854 AGCTTGACACACCTGCTTATCAACACAGCCTGCGGAGAGAGAGTGAATCTACTG 2913  
2638 TGTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2697  
2914 TGTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2973  
2698 CTGACACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2757  
2974 CTGACACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3033  
2758 CACCTGAGCCTTAGATGAACCTGCTGAGAGCAATAGGCTGAAGCTTCTGTGCGAGGTC 2817  
3034 CACCTGAGCCTTAGATGAACCTGCTGAGAGCAATAGGCTGAAGCTTCTGTGCGAGGTC 3093

QY 2818 ATGAGAAACCATCTTGTGATCTCCAGACCTGAGAGTGTGTAAGTGTATCTACCGCC 2877  
DB 3094 ATGAGAAACCATCTTGTGATCTCCAGACCTGAGAGTGTGTAAGTGTATCTACCGCC 3153  
QY 2878 GCGTCTGTGAGAGTCTGTCTGTGTGTATCTCGAGAGACAGACACCTGAAGAGCTGTAT 2937  
DB 3154 GCGTCTGTGAGAGTCTGTCTGTGTGTATCTCGAGAGACAGACACCTGAAGAGCTGTAT 3213  
QY 2938 CTCACGACATGCTCCCTGAGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2997  
DB 3214 CTCACGACATGCTCCCTGAGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3273  
QY 2998 AAGACAGT 3057  
DB 3274 AAGACAGT 3333  
QY 3058 GAGGACCTCTCCCTGAGT 3117  
DB 3334 GAGGACCTCTCCCTGAGT 3393  
QY 3118 AATACTTCACTCCCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3177  
DB 3394 AATACTTCACTCCCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3453  
QY 3178 AACTTACAGATTAATGAGT 3237  
DB 3454 AACTTACAGATTAATGAGT 3513  
QY 3238 GAGGAGTGCAGCTACTCAAGCCCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3297  
DB 3514 GAGGAGTGCAGCTACTCAAGCCCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3573  
QY 3298 GAGGATGACCGACAC 3312  
DB 3574 GAGGATGACCGGATAC 3588

RESULT 4  
US-10-399-443-23  
Sequence 23, Application US/10399443  
Publication No. US2004002869A1  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States of America, as Represented by the  
APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
APPLICANT: Health  
APPLICANT: Nelson, Lawrence M.  
APPLICANT: Tong, Zhi-Bin  
APPLICANT: Nelson, Lawrence  
APPLICANT: Zhi-Bin, Tong  
TITLE OF INVENTION: Human Gene Critical to Fertility  
FILE REFERENCE: 4239-64785  
CURRENT APPLICATION NUMBER: US/10/399,443  
CURRENT FILING DATE: 2003-04-16  
PRIOR APPLICATION NUMBER: 60/241,510  
PRIOR FILING DATE: 2000-10-18  
PRIOR APPLICATION NUMBER: PCT/US01/10981  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 23  
LENGTH: 3900  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(3600)  
OTHER INFORMATION:  
US-10-399-443-23

Query Match 78.5%; Score 3166.2; DB 17; Length 3900;  
Best Local Similarity 96.2%; Pred. No. 0;  
Matches 3304; Conservative 0; Mismatches 8; Indels 123; Gaps 1;

OY	1	NTGGAAGGAGCAAA	TGCTCTACCTTTTCAGCTACGGGCTGCAATGCTGTCTGATAG	60
Db	154	ATGGAAGGAGCAAA	TGCTCTACCTTTTCAGCTACGGGCTGCAATGCTGTCTGATAG	213
OY	61	CTAGACAAGGAAGATTT	CAGACATTCAGGAATTTACTTAAGAGAAATCTTCAGATCG	120
Db	214	CTAGACAAGGAAGATTT	CAGACATTCAGGAATTTACTTAAGAGAAATCTTCAGATCG	273
OY	121	ACCACATGCTCTATT	CCACAGTTGAAATCGAAGATGCGAACGTGGAAATGTCGGCACTC	180
Db	274	ACCACATGCTCTATT	CCACAGTTGAAATCGAAGATGCGAACGTGGAAATGTCGGCACTC	333
OY	181	CTCTTTCAGTAGTATT	TATGAGCAATCGCTGGCTGGGCTAGCTCATTAAGATCTTTGAA	240
Db	334	CTCTTTCAGTAGTATT	TATGAGCAATCGCTGGCTGGGCTAGCTCATTAAGATCTTTGAA	393
OY	241	AACATGAACCTGCGAA	CCCTCTCGGAGAAAGGCACGGATGCAATGAAAA-----	289
Db	394	AACATGAACCTGCGAA	CCCTCTCGGAGAAAGGCACGGATGCAATGAAAA-----	453
OY	290	-----	-----	289
Db	454	GAAAGTCTTGAAGCA	ACGATGCTGACCAAGGACCAAGCAAGAAAAAGTCCAGGAATT	513
OY	290	-----	-----AAATTCA	297
Db	514	TCACAAGCTGTGCA	CAAGATATGTCACAAGCTGCAAGACAAAAGAACAGAAAAATTTCA	573
OY	298	CAAGCTATGGAACA	GAAAGGTGCCCACAGACAGACAGAGAAGAAACAAGAAATTTCA	357
Db	574	CAAGCTATGGAACA	GAAAGGTGCCCACAGACAGACAGAGAAGAAACAAGAAATTTCA	633
OY	358	GCTATGGAACAAGAG	GTGCCACAGCAGACAGAGAAGAAACAAGACATGAGAGTAC	417
Db	634	GCTATGGAACAAGAG	GTGCCACAGCAGACAGAGAAGAAACAAGACATGAGAGTAC	693
OY	418	ACATGGAAGCTACA	AGATGTCAGTGTATGACCAAAATTCGTGAGAGAGAGATGTACTGT	477
Db	694	ACATGGAAGCTACA	AGATGTCAGTGTATGACCAAAATTCGTGAGAGAGAGATGTACTGT	753
OY	478	AGTTTGTAAAAACA	CTGCTGTGACCTGGGCGGAAATGCAAAAGTTGGCTGCTTTGAT	537
Db	754	AGTTTGTAAAAACA	CTGCTGTGACCTGGGCGGAAATGCAAAAGTTGGCTGCTTTGAT	813
OY	538	TCAGACCGGTGGGG	CTTCGCGCTTCGACCGGTGTTCTGCAACGAAATGCAAGATTGGG	597
Db	814	TCAGACCGGTGGGG	CTTCGCGCTTCGACCGGTGTTCTGCAACGAAATGCAAGATTGGG	873
OY	598	AAATGGGCTCTAG	CCAGAAAGATGCTGTCTGTCTGGGCGCAAGGTGACTCTACAGGGA	657
Db	874	AAATGGGCTCTAG	CCAGAAAGATGCTGTCTGTCTGGGCGCAAGGTGACTCTACAGGGA	933
OY	658	ATGTTCTCTCTAC	GTCTTCTCTCCCGTTTGAAGATGCAAGCGGAAAGAAAGAGCAGT	717
Db	934	ATGTTCTCTCTAC	GTCTTCTCTCCCGTTTGAAGATGCAAGCGGAAAGAAAGAGCAGT	993
OY	718	GTCAACAGATTCA	TCTCCAGGAGTGCGCAGACTCCAGGCTCCGGTACCGAAGATCATG	777
Db	994	GTCAACAGATTCA	TCTCCAGGAGTGCGCAGACTCCAGGCTCCGGTACCGAAGATCATG	1055
OY	778	TCCCGACCAAGAA	GGCTGTGTTCACTCATTTGACGGTTTCGATGACTTGGGCTCTGTCTC	837
Db	1054	TCCCGACCAAGAA	GGCTGTGTTCACTCATTTGACGGTTTCGATGACTTGGGCTCTGTCTC	1113
OY	838	AAACAATGACAA	AACTCTTGCAAAACCTGGGCTGAGAAAGCAAGCTTCGTTCACTCTATA	897
Db	1114	AAACAATGACAA	AACTCTTGCAAAACCTGGGCTGAGAAAGCAAGCTTCGTTCACTCTATA	1173
OY	898	CGCAGTCTGCTAG	AAGAGTCTCTGCTCCTGATGCTTCTCTGATGTCAACCGTCAGAGAC	957
Db	1174	CGCAGTCTGCTAG	AAGAGTCTCTGCTCCTGATGCTTCTCTGATGTCAACCGTCAGAGAC	1233
OY	958	GTGGGCAACAGAA	AGCTCAAGTCAAGAGTGTGTCTCCCGTTACTGTTAATGAGGA	1012

Dd	1234	GTGGGCACAGGAAGCTCAAGTCAAGAGTGTGTCTCCCGTTACCTGTTAGTAGAGA	1293
Qy	1018	ATCTCCGGGGAACAAGATTCACCTTCTCTTGGAGCGGGATTGGTAGCATCAGAAG	10777
Dd	1294	ATCTCCGGGGAACAAGATTCACCTTCTCTTGGAGCGGGATTGGTAGCATCAGAAG	1353
Qy	1078	ACACAAGGGTTGCGTGGATCATAGAACACCGTGAAGTGTCTCAACAGTCCAGGTGCC	1137
Dd	1354	ACACAAGGGTTGCGTGGATCATAGAACACCGTGAAGTGTCTCAACAGTCCAGGTGCC	1413
Qy	1138	GCCGTGGGCTCTCATCTGGGTGGCCCTGACGTGACGAGACGATGGTGGGGAAGGCTC	1197
Dd	1414	GCCGTGGGCTCTCATCTGGGTGGCCCTGACGTGACGAGACGATGGTGGGGAAGGCTC	1473
Qy	1198	GCCCCCTTCAACCAACCGCTCACAGGCGTGCACGCGCTTGTGTTCATCAGCTCAC	1257
Dd	1474	GCCCCCTTCAACCAACCGCTCACAGGCGTGCACGCGCTTGTGTTCATCAGCTCAC	1533
Qy	1258	CCTGAGGCGGTGTGTCCGCGCTGTCTCATCTGAGAGAAAGTGTCTGAGACGCTTC	1317
Dd	1534	CCTGAGGCGGTGTGTCCGCGCTGTCTCATCTGAGAGAAAGTGTCTGAGACGCTTC	1593
Qy	1318	TGCGGTATGGCTGTGAGGGAGTGTGGAATAGAAAGTCAAGTGTGATGTGAGACCTC	1377
Dd	1594	TGCGGTATGGCTGTGAGGGAGTGTGGAATAGAAAGTCAAGTGTGATGTGAGACCTC	1653
Qy	1378	ATGGTTCAGAGACTCGGGGAGTCTGAGCTCGTGCTGTGTTCAATGAACATCCTTCTC	1437
Dd	1654	ATGGTTCAGAGACTCGGGGAGTCTGAGCTCGTGCTGTGTTCAATGAACATCCTTCTC	1713
Qy	1438	CCAGACAGCCACTGTGAGAGTACTACACTTCTTCAACTCACTCAGTCTTCAGGACTTCTGT	1497
Dd	1714	CCAGACAGCCACTGTGAGAGTACTACACTTCTTCAACTCACTCAGGACTTCTGT	1773
Qy	1498	GCCCCCTTGTATCTACGTGTTAGAGGGCTGGAATATCAGCGAGTCTCTGCGCTGTAC	1557
Dd	1774	GCCCCCTTGTATCTACGTGTTAGAGGGCTGGAATATCAGCGAGTCTCTCTGCTGTAC	1833
Qy	1558	GTTGAGAAAGCAAAAGAGTCCATGAGCTTAAACAGGACGGCTTCATATCCATCGCTT	1617
Dd	1834	GTTGAGAAAGCAAAAGAGTCCATGAGCTTAAACAGGACGGCTTCATATCCATCGCTT	1893
Qy	1618	TGATTAAGAGCTTCTGTTTGGCTCGCTGTAGCGAAGCTTAAGAGGCGCACCTGAGAGTCTC	1677
Dd	1894	TGATTAAGAGCTTCTGTTTGGCTCGCTGTAGCGAAGCTTAAGAGGCGCACCTGAGAGTCTC	1953
Qy	1678	CTGCTGGGCTGTCCCGTTCCTCGGGGTGAAGAGAGCTTCTGCACTGGGTCTCTGTG	1737
Dd	1954	CTGCTGGGCTGTCCCGTTCCTCGGGGTGAAGAGAGCTTCTGCACTGGGTCTCTGTG	2013
Qy	1738	TTGGGTCAAGAGCTTAATGCCACACGCCAGAGACACCCCTGAGAGGCTTCACTGTCTT	1797
Dd	2014	TTGGGTCAAGAGCTTAATGCCACACGCCAGAGACACCCCTGAGAGGCTTCACTGTCTT	2073
Qy	1798	TTTGAAGCTCAAGCAAAAGCTTGTTCCTTGGCATTTAAACAGCTTCCAGAAAGTGTGG	1857
Dd	2074	TTTGAAGCTCAAGCAAAAGCTTGTTCCTTGGCATTTAAACAGCTTCCAGAAAGTGTGG	2133
Qy	1858	CTTCCGATTTAACGAAGCTTGAATGACATCTTCTTGTGCTTCCAGACTGTGCCG	1917
Dd	2134	CTTCCGATTTAACGAAGCTTGAATGACATCTTCTTGTGCTTCCAGACTGTGCCG	2193
Qy	1918	TATTTGGGAAATTTGGGTGATGTCAAAAGGATCTTCCCAAGATGAGTCCGCTGAG	1977
Dd	2194	TATTTGGGAAATTTGGGTGATGTCAAAAGGATCTTCCCAAGATGAGTCCGCTGAG	2253
Qy	1978	GCATGTCTGTGTCTCCCTATGAGATCGGAGTAAAGCCCTCATTTGAGAGCACTGGGAA	2037
Dd	2254	GCATGTCTGTGTCTCCCTATGAGATCGGAGTAAAGCCCTCATTTGAGAGCACTGGGAA	2113
Qy	2038	GATTTCTGTCTCATGTCTTGGCACCAACCCACCTGCGGACGTGACCTTGGGACGACG	2097



QY 290 -----AAATTCa 297  
Db 514 TCACAAGCTGTGCAACAAGATAGTGTCCACAGCTGCAGAGCAAAAAGAAACAGAAATTTTCA 573  
QY 298 CAAGCTATGAAACAAGAGGTGCCACAGCAGCAGAGCAAGAAACAAGAAATTTTCA 357  
Db 574 CAAGCTATGAAACAAGAGGTGCCACAGCAGCAGAGCAAGAAACAAGAAATTTTCA 633  
QY 358 GCTATGGAACAAGAGGTGCCACAGCAGCAGAGCAAGAAACAAGAAATTTTCA 417  
Db 634 GCTATGGAACAAGAGGTGCCACAGCAGCAGAGCAAGAAACAAGAAATTTTCA 693  
QY 418 ACATGGAGCTACAAGAGTCAAGTATGACCAATTCGCTGAGAGAGAGATGATAGTGTG 477  
Db 694 ACATGGAGCTACAAGAGTCAAGTATGACCAATTCGCTGAGAGAGAGATGATAGTGTG 753  
QY 478 AGTTTGAACAACATGCTGTGCTGTGCTGCGCGGAATGCAACGTTGGCTGTGTCTTTGAT 537  
Db 754 AGTTTGAACAACATGCTGTGCTGTGCTGCGCGGAATGCAACGTTGGCTGTGTCTTTGAT 813  
QY 538 TCAGACCGGTGGGGCTTCGCGCTGCGACGCGTGTCTGCAACGGAATGCAAGAAATTTGG 597  
Db 814 TCAGACCGGTGGGGCTTCGCGCTGCGACGCGTGTCTGCAACGGAATGCAAGAAATTTGG 873  
QY 598 AAATGGCTTAGCCAGAGAGATCGTGTGTGTGCTGTGCGCAAGGTGATCTTACAGGGA 657  
Db 874 AAATGGCTTAGCCAGAGAGATCGTGTGTGTGCTGTGCGCAAGGTGATCTTACAGGGA 933  
QY 658 ATGTTCTCTAAGTCTTCTTCTCCGCTTGAAGATGCAAGCGGAAGAAGAGACAGT 717  
Db 934 ATGTTCTCTAAGTCTTCTTCTCCGCTTGAAGATGCAAGCGGAAGAAGAGACAGT 993  
QY 718 GTCCAGAGTTCATCCAGGGAGTGGCCAGATCCCGAGGCTCCGCTGACGGAATGATG 777  
Db 994 GTCCAGAGTTCATCCAGGGAGTGGCCAGATCCCGAGGCTCCGCTGACGGAATGATG 1053  
QY 778 TCCCGACCAAGAGCTGTGTTCATCATTTGACGTTTCGATGACCTGTGCTGTCTC 837  
Db 1054 TCCCGACCAAGAGCTGTGTTCATCATTTGACGTTTCGATGACCTGTGCTGTCTC 1113  
QY 838 AACCAATGACAAAGCTCTGCAAGAGCTGGGCTGGAAGACAGCTCCGTTACCTCTATA 897  
Db 1114 AACCAATGACAAAGCTCTGCAAGAGCTGGGCTGGAAGACAGCTCCGTTACCTCTATA 1173  
QY 898 CGCAGCTGTGAGGAAGTCTGTCTCCCTGAGTCTTCCGATGATGATACCTGTGAGAG 957  
Db 1174 CGCAGCTGTGAGGAAGTCTGTCTCCCTGAGTCTTCCGATGATGATACCTGTGAGAG 1233  
QY 958 GTGGCACAAGAGCTCAAGTCAAGAGTCTGTCTCCCGTTACCTGTATGATTAGAGA 1017  
Db 1234 GTGGCACAAGAGCTCAAGTCAAGAGTCTGTCTCCCGTTACCTGTATGATTAGAGA 1293  
QY 1018 ATCTCCGGGAAACAAGAAATCCATTGCTCTTGAAGCGGGATTTGTGAGCATCAAG 1077  
Db 1294 ATCTCCGGGAAACAAGAAATCCATTGCTCTTGAAGCGGGATTTGTGAGCATCAAG 1353  
QY 1078 ACACAAGGTTGCGGTGATCAATGAACAACGTTGAGTCTGACCAAGTCCAGGTGCGC 1137  
Db 1354 ACACAAGGTTGCGGTGATCAATGAACAACGTTGAGTCTGACCAAGTCCAGGTGCGC 1413  
QY 1138 GCGGTGGCTCTCTCATCTGTGCTGTGCGCTGTGACGTGTGAGAGAGAGAGCGTC 1197  
Db 1414 GCGGTGGCTCTCTCATCTGTGCTGTGCGCTGTGACGTGTGAGAGAGAGAGCGTC 1473  
QY 1198 GCGCCCTTCAACAAACGCTCAAGAGCTGTGACGCGCTTTTGTGTTTCAATCACTCAC 1257  
Db 1474 GCGCCCTTCAACAAACGCTCAAGAGCTGTGACGCGCTTTTGTGTTTCAATCACTCAC 1533  
QY 1258 CCTGAGAGCTGTGCTGCGGTGTCTCAATCTGAGAGAAAGTTGTCTGAGAGCGTTC 1317  
Db 1534 CCTGAGAGCTGTGCTGCGGTGTCTCAATCTGAGAGAAAGTTGTCTGAGAGCGTTC 1593  
QY 1318 TGCCGTATGTGTGTGAGAGAGTGTGAATAGGAAGTCAAGTGTGATGTGACGACCTC 1377

Db 1594 TGCCGTATGTGTGTGAGAGAGTGTGAATAGGAATGCAAGTGTGATGTGAGACCTC 1653  
QY 1378 ATGCTTCAAGACTCTGGGAGTCTGAGCTCCGTCTCTGTTTCAATGAACATCTCTTC 1437  
Db 1654 ATGCTTCAAGACTCTGGGAGTCTGAGCTCCGTCTCTGTTTCAATGAACATCTCTTC 1713  
QY 1438 CCAGACAGCCACTGTGAGAGTACTACACCTTCTTCCAGCTCAGTCTCCAGACCTTCTGT 1497  
Db 1714 CCAGACAGCCACTGTGAGAGTACTACACCTTCTTCCAGCTCAGTCTCCAGACCTTCTGT 1773  
QY 1498 GCGCCCTTGTACTACGTTGATGAGAGGCTGTGAATTCAGACCAAGCTCTGTGCTGTAC 1557  
Db 1774 GCGCCCTTGTACTACGTTGATGAGAGGCTGTGAATTCAGACCAAGCTCTGTGCTGTAC 1833  
QY 1558 GTTGAAGACAAAGAGTCTCATGAGACTTTAAACAGCAGGCTTCCATATCCACTCCCT 1617  
Db 1834 GTTGAAGACAAAGAGTCTCATGAGACTTTAAACAGCAGGCTTCCATATCCACTCCCT 1893  
QY 1618 TGGATGAAGCGTTCTTGTGTGGCTGTGTGAGCAGAACGTTAAGAGAGCCACTGAGGTC 1677  
Db 1894 TGGATGAAGCGTTCTTGTGTGGCTGTGTGAGCAGAACGTTAAGAGAGCCACTGAGGTC 1953  
QY 1678 CTGCTGGCTGTCCCGTTCCCTGTGGGTGAAGCAGAGCTTGTGACTGTGCTCTGTG 1737  
Db 1954 CTGCTGGCTGTCCCGTTCCCTGTGGGTGAAGCAGAGCTTGTGACTGTGCTCTGTG 2013  
QY 1738 TTGGGTCAAGCAGCTTAATGCAACACCCAGAGACACCTGTGACGCTTCCACTGTCTT 1797  
Db 2014 TTGGGTCAAGCAGCTTAATGCAACACCCAGAGACACCTGTGACGCTTCCACTGTCTT 2073  
QY 1798 TTGAGACTCAAGACAAAGATTTGTTGCTGTGCAATTTAAACAGCTTCCAAAGAGTGTG 1857  
Db 2074 TTGAGACTCAAGACAAAGATTTGTTGCTGTGCAATTTAAACAGCTTCCAAAGAGTGTG 2133  
QY 1858 CTTCGATTTAACAGAACTGTGACTGTATAGCATTTCTTGTGCTTCCAGACTGTCTCG 1917  
Db 2134 CTTCGATTTAACAGAACTGTGACTGTATAGCATTTCTTGTGCTTCCAGACTGTCTCG 2193  
QY 1918 TATTTTGGGAAATTTCCGGGTGATGTCAAGAGATCTTCCCAAGATGATGTCCGCTGAG 1977  
Db 2194 TATTTTGGGAAATTTCCGGGTGATGTCAAGAGATCTTCCCAAGATGATGTCCGCTGAG 2253  
QY 1978 GCATGTCCGTGTGCTTCCCTCTATGATGAGTGGGGAATAAGACCTCATTTGAGAGCAGTGG 2037  
Db 2254 GCATGTCCGTGTGCTTCCCTCTATGATGAGTGGGGAATAAGACCTCATTTGAGAGCAGTGG 2133  
QY 2038 GATTTCTGCTCATGCTTTGGCACCACCCACACCTGTGCGGAGCTGTGACCTGGGACAGC 2097  
Db 2314 GATTTCTGCTCATGCTTTGGCACCACCCACACCTGTGCGGAGCTGTGACCTGGGACAGC 2373  
QY 2098 ATCTGACAGAGCGGGCTATGAACACCTGTGTGTCACAGTGTGAGATCCCACTGTGAG 2157  
Db 2374 ATCTGACAGAGCGGGCTATGAACACCTGTGTGTCACAGTGTGAGATCCCACTGTGAG 2433  
QY 2158 ATACAGACCTCGATGTTTGAATATGACAAGTTACCTCCGTGTGTGACAGCCTGTGAGA 2217  
Db 2434 ATACAGACCTCGATGTTTGAATATGACAAGTTACCTCCGTGTGTGACAGCCTGTGAGA 2493  
QY 2218 ATGCTCATGGCAACCGTTAACCTTAAGTCTCTCAATTTGGAGGACCCACTGAAGAA 2277  
Db 2494 ATGCTCATGGCAACCGTTAACCTTAAGTCTCTCAATTTGGAGGACCCACTGAAGAA 2553  
QY 2278 GAGGATGTAAGATGAGGTGTGAAGCTTAAACACCCAAATTTGTTGAGATCTTTG 2337  
Db 2554 GAGGATGTAAGATGAGGTGTGAAGCTTAAACACCCAAATTTGTTGAGATCTTTG 2613  
QY 2338 AGGCTGAGTTGCTGTGATTTGACCAATGCTGTTAACCTGAAGATCTCCAAATCTTACG 2397  
Db 2614 AGGCTGAGTTGCTGTGATTTGACCAATGCTGTTAACCTGAAGATCTCCAAATCTTACG 2673  
QY 2398 ACTTCCCCAGCCTGAATCTGTGAGCCTGTGCAAGAAACAAGTGTGACAGCAGAGATA 2457

Db	2674	ACCTCCCAAGCTGAATAATCTGTAGCCTGGAGAGAAACAAGTGAAGACCAAGGAGTA	2733
Qy	2458	ATGCTCTCAAGTATGCTTGAAGATCTCCAGTGCCTTCGAGAACTGATACTGAG	2517
Db	2734	ACGCTCTCAGTGAAGCCCTTGAGGGTCTCCAGTGCCTTCGAGAACTGATACTGAG	2793
Qy	2518	GACTGTGCATCAAGCAACGGGTTCAGAGTCTGGCTCAGGCCCTGTGAGCAACCG	2577
Db	2794	GACTGTGCATCAAGCAACGGGTTCAGAGTCTGGCTCAGGCCCTGTGAGCAACCG	2853
Qy	2578	AGCTTGACACACTGTGCCTATCCACAACAGCCTGGGGAACGAAGTGAATCTACTG	2637
Db	2854	AGCTTGACACACTGTGCCTATCCACAACAGCCTGGGGAACGAAGTGAATCTACTG	2913
Qy	2638	TGTGATCAATGAGGCTTCCCACTGATGTGACAGAGGCTGATGCTGAATCAGTCCAC	2697
Db	2914	TGTGATCAATGAGGCTTCCCACTGATGTGACAGAGGCTGATGCTGAATCAGTCCAC	2973
Qy	2698	CTGACAGGGCTGGCTGGTTCCTTGCACTTGGCTTAAGGGTAATCATGGCTGACG	2757
Db	2974	CTGACAGGGCTGGCTGGTTCCTTGCACTTGGCTTAAGGGTAATCATGGCTGACG	3033
Qy	2758	CACCTGACCTTGAAGTAAACCTGTGGAAGACAATGCGTGAAGCTTGTGCGAGGTC	2817
Db	3034	CACCTGACCTTGAAGTAAACCTGTGGAAGACAATGCGTGAAGCTTGTGCGAGGTC	3093
Qy	2818	ATGAGAGAACCATCTTGTCACTCCAGAGCTGAGTGGTAAAGTCACTCAACGCC	2877
Db	3094	ATGAGAGAACCATCTTGTCACTCCAGAGCTGAGTGGTAAAGTCACTCAACGCC	3153
Qy	2878	GCGTGTGTGAAGTCTGTCTGTGTATCTCGAGAGACACACCTGAAGAAGCTGGAT	2937
Db	3154	GCGTGTGTGAAGTCTGTCTGTGTATCTCGAGAGACACACCTGAAGAAGCTGGAT	3213
Qy	2938	CTACAGCAATGCGCTGGGTGACGAGTGGGGTGGCTGGCTGTGAGAGGACTGAAGCA	2997
Db	3214	CTACAGCAATGCGCTGGGTGACGAGTGGGGTGGCTGGCTGTGAGAGGACTGAAGCA	3273
Qy	2998	AAGAACAGTGTCTGACAGAGACTGGGTGGAAGGACATGACATGACTTCTGATGGCTGT	3057
Db	3274	AAGAACAGTGTCTGACAGAGACTGGGTGGAAGGACATGACATGACTTCTGATGGCTGT	3353
Qy	3058	GAGGCACTCTCCTTGGCCCTTTCCTGCAACCGGCACTGACCAAGTCTAAACCTGGTCAG	3117
Db	3334	GAGGCACTCTCCTTGGCCCTTTCCTGCAACCGGCACTGACCAAGTCTAAACCTGGTCAG	3393
Qy	3118	AATAAATTCAATCCCAAAAGAAATGAAGTGTGTGGGCTTGTGCTTCCACGTCT	3177
Db	3394	AATAAATTCAATCCCAAAAGAAATGAAGTGTGTGGGCTTGTGCTTCCACGTCT	3453
Qy	3178	AACCTTACAGATTAATTGGGCTGTGAAATAGGAGTACCTGTGCAAAATAAGGAAGTCTGT	3257
Db	3454	AACCTTACAGATTAATTGGGCTGTGAAATAGGAGTACCTGTGCAAAATAAGGAAGTCTGT	3513
Qy	3238	GAGGAAGTGACGTAATCAAGCCCCGAGTCTGAATTGACGGTAGTGGCAATCTTTTGAT	3297
Db	3514	GAGGAAGTGACGTAATCAAGCCCCGAGTCTGAATTGACGGTAGTGGCAATCTTTTGAT	3573
Qy	3298	GAAATGACCGACAC 3312	
Db	3574	GAAATGACCGGTAC 3588	

RESULT 6  
US-10-216-645-3

Sequence 3, Application US/10216645  
Publication No. US20030125282A1  
GENERAL INFORMATION:  
APPLICANT: WEISS, BERTTRAM  
APPLICANT: LESSL, MONIKA  
APPLICANT: PETERS-KOTTIG, MICHAEL  
APPLICANT: BECKMANN, GORG  
TITLE OF INVENTION: HUMAN MATR PROTEINS

```

; FILE REFERENCE: SCH-1910
; CURRENT APPLICATION NUMBER: US/10/216,645
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3830
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-216-645-3

```

Query Match	77.8%	Score 3139;	DB 15;	Length 3830;
Best Local Similarity	96.0%;	Pred. No. 0;		
Matches 3282; Conservative	0;	Mismatches 30;	Indels 105;	Gaps 2;

0y 1 ATGAGAGAGACAATCGCTCACTTTTCAGCTACGGCTGCATGAGTCTTATAG 60

Db 1 ATGAGAGAGACAATCGCTCACTTTTCAGCTACGGCTGCATGAGTCTTATAG 60

61 CTGACCAAGGAAGAAATTCAGACATTCAAGAAATCTTCAAGATCG 120  
61 CTGACCAAGGAAGAAATTCAGACATTCTTAAAGAAATCTTCAAGATCG 120

121 ACCACATCTCTATTCCACAGTTGAAATCGAATGCCAAGCGAATGTCGTGGACATC 180

Db 181 CTTTGACATGAGATTATGAGACATCGTGGCCTGGGCTACGTCATTATGACATCTTTGAA 240

Db 241 AACATGAACCTGGCAACCTCTTCGGAGAGGCGACGGGATGACATGAANAATTCCACGAGA 300

301 GATCTGAAGCAACGATCACTGACCAAGGACCAAGCAAGGAAAGTCCAGAAATAA 360

[illegible]

Db 421 AAGTATGTTGAATTCATCTCTTTTGGCAGAAATTTCAACAAGCTATGGAACAAGAAAGT 480

481 GCACACAGCAGAGACGAAAGACATGGAGTGACCATGGGACTTCMAAGGT 540

496 GCTGACTGGCCGGAAATGCAAAAGTTGGCTGGTGGCTTTTGATTACAGACCGGTCGGGGCTTTC 555  
 Db 541 CACGTGATGACCAATTCGCTGAGAGAGAGATGTACCTCGTAGTGTGTTGAAAAACAATCTGT 600

Dd 601 GCTGACCTGGCCGGAAATGCAAACGTTGCTGGTCTTTTGAATTCCAGACCGGTGGGCGCTTC 660

Dy 556 CGGCTTCGCACGGTGTTCTGCACGGAAAGTCAGGAATTGGGAAATCGGCTTAGCCAGA 615

Dd		661 CGGGCTCGCACGSGTGTCTCTGCACGGAAGTCAGAATTGGGAAATCGGCTCTAGCCACA	720
Gy	616 AGAGTCGTCGTCGTCGGGCCGAAGTGACCTCAACAGGAAATGTTCTCCTCATOGTCTTC	675	

D6	721	AGAGATCGTGTGTCGTCGGCGCAAGGTGGACATCTACCAAGGAGATGTTCTCTCAAGTCTTC	780
Oy	676	TTCCTCCCCGTGAGAGATGCAAGCGGAAAGAGAGACAGTCAAGAGTTCATCTCC	735

[illegible]

QY 736 AGGAGTGGCCAGACTCCCAAGGCTCCGGTGAAGGATCATGTCCCGACAGAAAAGCGTG 795  
Db 841 AGGAGTGGCCAGACTCCCAAGGCTCCGGTGAAGGATCATGTCCCGACAGAAAAGCGTG 900  
QY 796 TTGTTCAATGATGACGTTTGATGACCTGGGCTCTGTCTTCAACATGACACAAAGCTC 855  
Db 901 TTGTTCAATGATGACGTTTGATGACCTGGGCTCTGTCTTCAACAAATGACACAAAGCTC 960  
QY 856 TGCAGAGCTGGGCTGAGAGAGCCTCGTTCAACCTCATACGACGCTGTGAGAG 915  
Db 961 TGCAGAGCTGGGCTGAGAGAGCCTCGTTCAACCTCATACGACGCTGTGAGAG 1020  
QY 916 GTCCGTCTCCCTGATCCTTCTGATCGTCAACCGTCAGAGAGCTGGGACAGAGAGCTC 975  
Db 1021 GTCCGTCTCCCTGATCCTTCTGATCGTCAACCGTCAGAGAGCTGGGACAGAGAGCTC 1080  
QY 976 AAGTCAGAGTGTGTCTTCCCGTTACCTGTTAGTAGAGAAATCTCCGGGAAACAAAGA 1035  
Db 1081 AAGTCAGAGTGTGTCTTCCCGTTACCTGTTAGTAGAGAAATCTCCGGGAAACAAAGA 1140  
QY 1036 ATCCACTTGTCTTGAAGCGGGGATGTGTAGCATCAAGAGACACAGGGGTGTGCG 1095  
Db 1141 ATCCACTTGTCTTGAAGCGGGGATGTGTAGCATCAAGAGACACAGGGGTGTGCG 1200  
QY 1096 ATCATGAACAAACCGTGAAGCTGTCCAGACAGTGCACGCGCGCGTGGGCTCTTCATC 1155  
Db 1201 ATCATGAACAAACCGTGAAGCTGTCCAGACAGTGCACGCGCGCGTGGGCTCTTCATC 1260  
QY 1156 TGCCTGCTTGCACCTGACAGAGCGTGTGGGGAGAGCGTGCCTTCAACCAACG 1215  
Db 1261 TGCCTGCTTGCACCTGACAGAGCGTGTGGGGAGAGCGTGCCTTCAACCAACG 1320  
QY 1216 CTCACAGGCTGACAGCGGCTTTGTGTTCATCAAGCTCACCCCTGAGGGTGTCCGG 1275  
Db 1321 CTCACAGGCTGACAGCGGCTTTGTGTTCATCAAGCTCACCCCTGAGGGTGTCCGG 1380  
QY 1276 CGCTGTCTCAATCTTGAAGAAAGATGTCTTGAAGCGCTTCTGCGTATGAGTGTGAG 1335  
Db 1381 CGCTGTCTCAATCTTGAAGAAAGATGTCTTGAAGCGCTTCTGCGTATGAGTGTGAG 1440  
QY 1336 GGAAGTGTGAATGAGAACTCAGTGTGTTGATGTGACGACCTTCATGTTCAAGAGCTCGGG 1395  
Db 1441 GGAAGTGTGAATGAGAACTCAGTGTGTTGATGTGACGACCTTCATGTTCAAGAGCTCGGG 1500  
QY 1396 GAGTGTGAGCTCCGCTCTGTTCACATGAACATCTTCCCAAGACAGGCACTGTAG 1455  
Db 1501 GAGTGTGAGCTCCGCTCTGTTCACATGAACATCTTCCCAAGACAGGCACTGTAG 1560  
QY 1456 GAGTGTACACCTTCTTCAACCTCAGTCTCCAGACTTCTGTGCGGCTTGTACTACGTG 1515  
Db 1561 GAGTGTACACCTTCTTCAACCTCAGTCTCCAGACTTCTGTGCGGCTTGTACTACGTG 1620  
QY 1516 TTAGAGGCTTGAATGAGCAGCTCTTCCCTCTGTACGTTGAGAGACAAAGAG 1575  
Db 1621 TTAGAGGCTTGAATGAGCAGCTCTTCCCTCTGTACGTTGAGAGACAAAGAG 1680  
QY 1576 TCCATGAGCTTAAACAAGCAGGCTTCAATACCATGCTTGTGAGATGAAGCTTCTTG 1635  
Db 1681 TCCATGAGCTTAAACAAGCAGGCTTCAATACCATGCTTGTGAGATGAAGCTTCTTG 1740  
QY 1636 TTTGGCTCTGTGAGCGAAGACGTATGAGAGGCACTGAGAGTCTCTGGGCTGTCCGTT 1695  
Db 1741 TTTGGCTCTGTGAGCGAAGACGTATGAGAGGCACTGAGAGTCTCTGGGCTGTCCGTT 1800  
QY 1696 CCCCTGGGGTGAACAGAAAGCTTCTGCACTGGGTCTCTCTGTGGGTCAAGACCTTAAT 1755  
Db 1801 CCCCTGGGGTGAACAGAAAGCTTCTGCACTGGGTCTCTCTGTGGGTCAAGACCTTAAT 1860  
QY 1756 GGCACCAACCCAGAGAGACCTGTGACGCTTCCCATCTTTCGAGACTCAAGACAA 1815  
Db 1861 GGCACCAACCCAGAGAGACCTGTGACGCTTCCCATCTTTCGAGACTCAAGACAA 1920  
QY 1816 GAGTTTGTTCGCTTGGCATTAACAGCTTCCAAAGAGTGTGGCTTCCGATTAAACAGAAC 1875

Db 1921 GAGTTTGTTCGCTTGGCATTAACAGCTTCCAAAGAGTGTGGCTTCCGATTAAACAGAAC 1980  
QY 1876 CTGACCTTGAATAGATCTTCTTGTGCTTCCAGACACTGTCCGTAATTTGGCGAAATTCGG 1935  
Db 1981 CTGACCTTGAATAGATCTTCTTGTGCTTCCAGACACTGTCCGTAATTTGGCGAAATTCGG 2040  
QY 1936 GTGATGTCAAAAGGATCTTCCCAAGAGATGATGCTCCGCTGAGGATGCTGTGGTCCCT 1995  
Db 2041 GTGATGTCAAAAGGATCTTCCCAAGAGATGATGCTCCGCTGAGGATGCTGTGGTCCCT 2100  
QY 1996 CTATGATGCGGATTAAGACCTTCATTGAGAGAGAGTGGAAAGATTTCTGTCCATGCTT 2055  
Db 2101 CTATGATGCGGATTAAGACCTTCATTGAGAGAGAGTGGAAAGATTTCTGTCCATGCTT 2160  
QY 2056 GGCACCAACCCAGACCTGCGGACGCTGACCTGGGACAGACATCTTCAAGAGCGGGCC 2115  
Db 2161 GGCACCAACCCAGACCTGCGGACGCTGACCTGGGACAGACATCTTCAAGAGCGGGCC 2220  
QY 2116 ATGAAGACCCGTGTGTGCAAGCTGAGGATCCCACTGCAAGATPACAGACCTGATGTTT 2175  
Db 2221 ATGAAGACCCGTGTGTGCAAGCTGAGGATCCCACTGCAAGATPACAGACCTGATGTTT 2280  
QY 2176 AGAATGACAGATTACCCCTGTGTGCAAGACCTTGAAGATGTCATGAGCCAAACGT 2235  
Db 2281 AGAATGACAGATTACCCCTGTGTGCAAGACCTTGAAGATGTCATGAGCCAAACGT 2340  
QY 2236 AACCTAAGATCTTCAACTTGGAGGACCACTTGAAGAGAGATGTAAGATGGCG 2295  
Db 2341 AACCTAAGATCTTCAACTTGGAGGACCACTTGAAGAGAGATGTAAGATGGCG 2400  
QY 2296 TGTGAAGCTTAAACACCCAAATGTTTGTGAGATCTTGAAGGCTGATTTGCTGTGGA 2355  
Db 2401 TGTGAAGCTTAAACACCCAAATGTTTGTGAGATCTTGAAGGCTGATTTGCTGTGGA 2460  
QY 2356 TTGACCAATGCTTAACTTGAAGATCTCCAAATCTTACAGACTCCCGACGCTGAAA 2415  
Db 2461 TTGACCAATGCTTAACTTGAAGATCTCCAAATCTTACAGACTCCCGACGCTGAAA 2520  
QY 2416 TCTCTGAGCTGCGAGAGAAACAGATGACAGACCAAGGAGATTAATGCTTCAAGATGCC 2475  
Db 2521 TCTCTGAGCTGCGAGAGAAACAGATGACAGACCAAGGAGATTAATGCTTCAAGATGCC 2580  
QY 2476 TTGAGAGTCTCCAGTGGCGCTTCAAGAGCTGATTACTGAGAGACTGTGGACTACAGGC 2535  
Db 2581 TTGAGAGTCTCCAGTGGCGCTTCAAGAGCTGATTACTGAGAGACTGTGGACTACAGGC 2640  
QY 2536 ACGGTTGCGAGAGTGGCTTCAAGCCTGTGTAGCAACCGAGCTTGAACAACCTGTGC 2595  
Db 2641 ACGGTTGCGAGAGTGGCTTCAAGCCTGTGTAGCAACCGAGCTTGAACAACCTGTGC 2700  
QY 2596 CTATCCAAACAACCTGCGGAGACAGAGGTAAATCTATCTGTGTCAATGATGAGCTT 2655  
Db 2701 CTATCCAAACAACCTGCGGAGACAGAGGTAAATCTATCTGTGTCAATGATGAGCTT 2760  
QY 2656 CCCCACTGTATGTGAGAGGCTGATGTGAATGAGGCAACCTGAGACACGCGTGGGCTGT 2715  
Db 2761 CCCCACTGTATGTGAGAGGCTGATGTGAATGAGGCAACCTGAGACACGCGTGGGCTGT 2820  
QY 2716 GGTTTTCTTGGACTTGGCTTATGGGTAACTCATGAGCTGACGACTGAGGCTTAGCATG 2775  
Db 2821 GGTTTTCTTGGACTTGGCTTATGGGTAACTCATGAGCTGACGACTGAGGCTTAGCATG 2880  
QY 2776 AACCTGTGAAAGACATGCGGTGTAAGCTTGTGTGCGAGGTCAATGAAGAACATCTTGT 2835  
Db 2881 AACCTGTGAAAGACATGCGGTGTAAGCTTGTGTGCGAGGTCAATGAAGAACATCTTGT 2940  
QY 2836 CATCTCAGAGACCTGAGATTGTAAAGTGCATCTCACCGCGGTGTGTGAGAGCTG 2895  
Db 2941 CATCTCAGAGACCTGAGATTGTAAAGTGCATCTCACCGCGGTGTGTGAGAGCTG 3000  
QY 2896 TCTGTGTGATCTTGAAGAGACACCTGAAGAGCTGTGATCTCAAGCAATGCTGCTG 2955



Db 3001 TCCTGTGTGATCTCGAGGACAGACACTGAAGAGCCTGATCTCAACGACATGCCCC 3060  
Qy 2956 GGTGACGGTGGGTTGCTGCGCTGTGCGGAGGAGCTGAAGAAAGAAAGCTGTTCTGAG 3015  
Db 3061 GGTGACGGTGGGTTGCTGCGCTGTGCGGAGGAGCTGAAGAAAGAAAGCTGTTCTGAG 3120  
Qy 3016 AGACTCGGGTTGAAGGACATGTGAGACTGATCTGATGTTGTGAGGACATCTCCCTGGCC 3075  
Db 3121 AGACTCGGGTTGAAGGACATGTGAGACTGATCTGATGTTGTGAGGACATCTCCCTGGCC 3180  
Qy 3076 CTTTCTCGACACCGGATCTGACCACTTAAACCTGTGACAGAAATTAATCTTCACTCCCAA 3135  
Db 3181 CTTTCTCGACACCGGATCTGACCACTTAAACCTGTGACAGAAATTAATCTTCACTCCCAA 3240  
Qy 3136 GGAATGATGAGCTGTGTTGCGGCTTTGCTGCTCCAGCTCTTAATCTTACATTAATGGG 3195  
Db 3241 GGAATGATGAGCTGTGTTGCGGCTTTGCTGCTCCAGCTCTTAATCTTACATTAATGGG 3300  
Qy 3196 CTGTGAAATGAGCAGTACCTGTGCAAAATGAAGAGCTGTGAGGAGTGAAGCTGCTC 3255  
Db 3301 CTGTGAAATGAGCAGTACCTGTGCAAAATGAAGAGCTGTGAGGAGTGAAGCTGCTC 3360  
Qy 3256 AAGCCCGAGTGTGATTTGACCGTATGTTGGCATTTCTTTGATGAAGATGACCGAC 3312  
Db 3361 AAGCCCGAGTGTGATTTGACCGTATGTTGGCATTTCTTTGATGAAGATGACCGGTA 3417

RESULT 7  
US-10-416-642-3  
; Sequence 3, Application US/10416642  
; Publication No. US20040043452A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: RAMKUMAR, Jayalaxmi  
; APPLICANT: ARVIZU, Chandra  
; TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0842 PCT  
; CURRENT APPLICATION NUMBER: US/10/416,642  
; CURRENT FILING DATE: 2003-05-13  
; PRIOR APPLICATION NUMBER: 60/249,407  
; PRIOR FILING DATE: 2000-11-15  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PERL Program  
; SEQ ID NO: 3  
; LENGTH: 3489  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040043452A1 7474830CB1  
US-10-416-642-3

Query Match 77.7%; Score 3133.6; DB 18; Length 3489;  
Best Local Similarity 95.2%; Pred. No. 0;  
Matches 3308; Conservative 0; Mismatches 4; Indels 162; Gaps 1;

Qy 1 ATGAAGAGAGCAAAATCGCTACCTTTTCAGCTACCGGCTGCAATGTGTCTATGAG 60  
Db 1 ATGAAGAGAGCAAAATCGCTACCTTTTCAGCTACCGGCTGCAATGTGTCTATGAG 60  
Qy 61 CTAGCAAGAGAAATTTTCAAGCATTTCAAGAAATTAATTAAGAAATTTTCAAGATCG 120  
Db 61 CTAGCAAGAGAAATTTTCAAGCATTTCAAGAAATTAATTAAGAAATTTTCAAGATCG 120  
Qy 121 ACCAATGCTCTATTTCCACAGTTTGAATTCAGAAATGCAAGTGAATGTCTGACATC 180  
Db 121 ACCAATGCTCTATTTCCACAGTTTGAATTCAGAAATGCAAGTGAATGTCTGACATC 180  
Qy 181 CTCTTGATGATATTAATGAGCATCGCTGGCTGAGCTGATTAATGATGATCTTTGAA 240  
Db 181 CTCTTGATGATATTAATGAGCATCGCTGGCTGAGCTGATTAATGATGATCTTTGAA 240  
Qy 241 AAGATGAACCTGCAACCTCTCGAGAGAGGACCGGATGATGATGAAAA----- 289

Db 241 AAGATGAACCTGCAACCTCTCGAGAGAGGACCGGATGATGATGAAAAATTTCAACAGAA 300  
Qy 290 ----- 289  
Db 301 GATCTGAAGAACGATGACTGACCAAGAACCAAGAAAGAAAGTCCAGAAAAATAA 360  
Qy 290 ----- 289  
Db 361 TATGGCATGACTAAGCTTATCTTGGGGTGTCTGACATCTGACTCGAATTAATAAC 420  
Qy 290 ----- AAATTTCAAGCTATGAAACAAAGAT 318  
Db 421 AAGTATGTTGAATTCATTTCTTTTTCAGAAAAATTTTCAAGCTATGAAACAAAGAT 480  
Qy 319 GCCACAGCAGACAGACAGAAAGAAACAAAGAAATTTTCAAGCTATGAAACAAAGATGTC 378  
Db 481 GCCACAGCAGACAGACAGAAAGAAACAAAGAAATTTTCAAGCTATGAAACAAAGATGTC 540  
Qy 379 AAGAGAGAGAGACAGAAAGAAACAAAGACATGAGGTGACATGAGGACTTACAGAGTAC 438  
Db 541 AAGAGAGAGAGACAGAAAGAAACAAAGACATGAGGTGACATGAGGACTTACAGAGTAC 600  
Qy 439 GTGATGACCAAAATTTGCTGAGAGAGAGATGATGATGATGATGATGATGATGATGATG 498  
Db 601 GTGATGACCAAAATTTGCTGAGAGAGAGATGATGATGATGATGATGATGATGATGATG 660  
Qy 499 GACTGGCCGGAATGCAAACTGTTGCTGTTGATGATGATGATGATGATGATGATGATG 558  
Db 661 GACTGGCCGGAATGCAAACTGTTGCTGTTGATGATGATGATGATGATGATGATGATG 720  
Qy 559 CTTGCGACAGTGTGTTGCTGACAGAAAGTCAAGAAATTTGGAATTTGCTGACAGAAAG 618  
Db 721 CTTGCGACAGTGTGTTGCTGACAGAAAGTCAAGAAATTTGGAATTTGCTGACAGAAAG 780  
Qy 619 ATCGGCTGTGCTGGGCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 678  
Db 781 ATCGGCTGTGCTGGGCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 840  
Qy 679 CTCCCGGTTAAGAGATGACAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 738  
Db 841 CTCCCGGTTAAGAGATGACAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
Qy 739 GATGAGCCAGACTCCAGGCTCCGCTGACAGAGATCATGTCCTCCAGCCAGAAAGCTGTTG 798  
Db 901 GATGAGCCAGACTCCAGGCTCCGCTGACAGAGATCATGTCCTCCAGCCAGAAAGCTGTTG 960  
Qy 799 TTTCAATCATTTGACGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 858  
Db 961 TTTCAATCATTTGACGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
Qy 859 AAGAGCTGGGCTGAGAGACAGCTCCGTTTCAACCTCATACAGCTGCTGAGAGAGTTC 918  
Db 1021 AAGAGCTGGGCTGAGAGACAGCTCCGTTTCAACCTCATACAGCTGCTGAGAGAGTTC 1080  
Qy 919 CTGCTCCCTGAGTCTTCTCTATGCTCACTGCTGAGAGAGAGAGAGAGAGAGAGTCAAG 978  
Db 1081 CTGCTCCCTGAGTCTTCTCTATGCTCACTGCTGAGAGAGAGAGAGAGAGAGAGTCAAG 1140  
Qy 979 TCAGAGGTGCTGCTCCCGCTTACCTGTTAGTTAGAGAAATCTCCGGGAAACAAAGATTC 1038  
Db 1141 TCAGAGGTGCTGCTCCCGCTTACCTGTTAGTTAGAGAAATCTCCGGGAAACAAAGATTC 1200  
Qy 1039 CACTTGTCTCTTGAAGCGGAGATTGTGAGCATGAGAGACAAAGAGTTCGTCATCTGC 1098  
Db 1201 CACTTGTCTCTTGAAGCGGAGATTGTGAGCATGAGAGACAAAGAGTTCGTCATCTGC 1260  
Qy 1099 ATGAACAAAGCTGAGCTGCTGACCAAGTCCAGAGTCCCGCTGAGGCTCTCTCATCTGC 1158  
Db 1261 ATGAACAAAGCTGAGCTGCTGACCAAGTCCAGAGTCCCGCTGAGGCTCTCTCATCTGC 1320  
Qy 1159 GTGGCCCTGAGCTGAGAGAGTGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1218

Db 1321 GTGGCCCTGCAGCTGCAGAGACGTGTGGGGAGAGCGTCGCCCTTCAACCAACGCTC 1380  
Qy 1219 ACAGGCTGCAGCGCGCTTTGTGTTCATGAGCTCAACCTCCAGAGGCTGTCCGGGC 1278  
Db 1381 ACAGGCTGCAGCGCGCTTTGTGTTCATGAGCTCAACCTCCAGAGGCTGTCCGGGC 1440  
Qy 1279 TGTCTCAATCTGCAGGAAGAGTTGTCTTGAAGCGCTTCTCCGTAATGCTGTGGAGGA 1338  
Db 1441 TGTCTCAATCTGCAGGAAGAGTTGTCTTGAAGCGCTTCTCCGTAATGCTGTGGAGGA 1500  
Qy 1339 GTGTGAATAGAAAGTCACTGTGTGTGATGTCACCTTCATGTGTTCAAGACTGGGGAG 1398  
Db 1501 GTGTGAATAGAAAGTCACTGTGTGTGATGTCACCTTCATGTGTTCAAGACTGGGGAG 1560  
Qy 1399 TCTGAGCTCCGAGCTGTGTTCAATGAACATCCCTTCCAGACAGCCACTGTGAGAG 1458  
Db 1581 TCTGAGCTCCGAGCTGTGTTCAATGAACATCCCTTCCAGACAGCCACTGTGAGAG 1620  
Qy 1459 TACTACACCTTCTTCAACCTGAGTCTCCAGACCTTGTGCGCGCTTGTACTAGCTGTTA 1518  
Db 1621 TACTACACCTTCTTCAACCTGAGTCTCCAGACCTTGTGCGCGCTTGTACTAGCTGTTA 1680  
Qy 1519 GAGGCGCTGCAGAAATGAGAGCAGCTCTGCGCTCTGTACGTTGAGAGACAAAGAGTCC 1578  
Db 1681 GAGGCGCTGCAGAAATGAGAGCAGCTCTGCGCTCTGTACGTTGAGAGACAAAGAGTCC 1740  
Qy 1579 ATGAGCTTAAACAGGAGGCTTCCATATCACTGCTTGGATGAGAGCTTCTGTGTT 1638  
Db 1741 ATGAGCTTAAACAGGAGGCTTCCATATCACTGCTTGGATGAGAGCTTCTGTGTT 1800  
Qy 1639 GGCCTCTGAGCGAGAGACGTAAAGAGGCGCACTGAGAGTCTGTGCGGCTGTCCGTTCC 1698  
Db 1801 GGCCTCTGAGCGAGAGACGTAAAGAGGCGCACTGAGAGTCTGTGCGGCTGTCCGTTCC 1860  
Qy 1699 CTGGGCGTGAAGCAGAACTTCTGCACTGGGCTCTCTGTGTGGGCTCAGCAACCTAATGCC 1758  
Db 1861 CTGGGCGTGAAGCAGAACTTCTGCACTGGGCTCTCTGTGTGGGCTCAGCAACCTAATGCC 1920  
Qy 1759 ACCACCCAGAGAGACCGCTGAGCGCTTCCACTGTCTTTGAGAGCTCAAGACAAAGAG 1818  
Db 1921 ACCACCCAGAGAGACCGCTGAGCGCTTCCACTGTCTTTGAGAGCTCAAGACAAAGAG 1980  
Qy 1819 TTTGTTGCTTGGCAATTAAGCTTCCAAAGAGTGTGGCTTCCGATTAAACAGAACCTG 1878  
Db 1981 TTTGTTGCTTGGCAATTAAGCTTCCAAAGAGTGTGGCTTCCGATTAAACAGAACCTG 2040  
Qy 1879 GACTGTATAGCATCTTCTTCTGCTCTCAAGCACTGTCCGTATTTGCGGAAATTCGGGTG 1938  
Db 2041 GACTGTATAGCATCTTCTTCTGCTCTCAAGCACTGTCCGTATTTGCGGAAATTCGGGTG 2100  
Qy 1939 GATGTCAAAAGGATCTTCCCAAGAGATGTCGCGTGAAGCATGTCTGTGCTCCCTCTA 1998  
Db 2101 GATGTCAAAAGGATCTTCCCAAGAGATGTCGCGTGAAGCATGTCTGTGCTCCCTCTA 2160  
Qy 1999 TGGATGCGGATTAAGACCTCAATTGAGAGAGTGTGGAAATTTCTGCTCAATGTGGC 2058  
Db 2161 TGGATGCGGATTAAGACCTCAATTGAGAGAGTGTGGAAATTTCTGCTCAATGTGGC 2220  
Qy 2059 ACCCAACCAACCTGCGGAGCTGAGACCTGGGAGAGAGATCTGACAGAGCGGCGCATG 2118  
Db 2221 ACCCAACCAACCTGCGGAGCTGAGACCTGGGAGAGAGATCTGACAGAGCGGCGCATG 2280  
Qy 2119 AAGACCCGTGTGCCAAGCTGAGAGCATCCACCTGCAAGATACAGACCTGATGTTTGA 2178  
Db 2281 AAGACCCGTGTGCCAAGCTGAGAGCATCCACCTGCAAGATACAGACCTGATGTTTGA 2340  
Qy 2179 AATGACAGATTAACCTCTGTGTGAGAGACCTCTGAGAAATCTGTATGGCCACCTATAC 2238  
Db 2341 AATGACAGATTAACCTCTGTGTGAGAGACCTCTGAGAAATCTGTATGGCCACCTATAC 2400  
Qy 2239 CTAAAGATCCCTCAACTTGGAGAGCAACCACTGAAGAAAGAGATGAAGATGCGGTGT 2298  
Db 2401 CTAAAGATCCCTCAACTTGGAGAGCAACCACTGAAGAAAGAGATGAAGATGCGGTGT 2460

Qy 2299 GAAAGCTTAAAAACCCCAAAATGTTTGTGAGAGTCTTGAAGGCTGATGTCGTGATG 2358  
Db 2461 GAAAGCTTAAAAACCCCAAAATGTTTGTGAGAGTCTTGAAGGCTGATGTCGTGATG 2520  
Qy 2359 ACCCATGCTGTATCTGAAAGATCTCCAAATCTTACGACCTTCCAGCTGAAATCT 2418  
Db 2521 ACCCATGCTGTATCTGAAAGATCTCCAAATCTTACGACCTTCCAGCTGAAATCT 2580  
Qy 2419 CTGAGCTGCGAAGAAACAAGGTGACAGACAGGAGAAATATGCTCTCAGTGAAGCTTG 2478  
Db 2581 CTGAGCTGCGAAGAAACAAGGTGACAGACAGGAGAAATATGCTCTCAGTGAAGCTTG 2640  
Qy 2479 AGAGTCTCCAGTGCAGGCTCTGAGAACTGATATCTGAGAGACTGTGACATACAGCACG 2538  
Db 2641 AGAGTCTCCAGTGCAGGCTCTGAGAACTGATATCTGAGAGACTGTGACATACAGCACG 2700  
Qy 2539 GGTTCGAGAGTCTGAGCTTCAAGCTTGTGACGACCGGAGCTTGAACACCTGTGCTTA 2598  
Db 2701 GGTTCGAGAGTCTGAGCTTCAAGCTTGTGACGACCGGAGCTTGAACACCTGTGCTTA 2760  
Qy 2599 TCCAAACAAGCTTGGGGAACGAAAGGTAAATCTATGTGTGATGCAATGAGGCTTCCC 2658  
Db 2761 TCCAAACAAGCTTGGGGAACGAAAGGTAAATCTATGTGTGATGCAATGAGGCTTCCC 2820  
Qy 2659 CACTGTAGTCTGCAGAGGCTGATGCTGAATCAGTGCACCTGAGACAGGCTGTGTGT 2718  
Db 2821 CACTGTAGTCTGCAGAGGCTGATGCTGAATCAGTGCACCTGAGACAGGCTGTGTGT 2880  
Qy 2719 TTTCTTGCATCTGGCTTATGAGTAACTCATGCTGACGCACTGAGCCTTAAAGATGAC 2778  
Db 2881 TTTCTTGCATCTGGCTTATGAGTAACTCATGCTGACGCACTGAGCCTTAAAGATGAC 2940  
Qy 2779 CCTGTGAAGAACAATGCGGTGAAGCTTCTGTGCGAGGTCAATGAGAGAACATCTGTGCAT 2838  
Db 2941 CCTGTGAAGAACAATGCGGTGAAGCTTCTGTGCGAGGTCAATGAGAGAACATCTGTGCAT 3000  
Qy 2839 CTCAGAGACCTGAGATTGTAAAGTGCATCTCAACCCCGGCTGTGAGAGTCTGTGCC 2898  
Db 3001 CTCAGAGACCTGAGATTGTAAAGTGCATCTCAACCCCGGCTGTGAGAGTCTGTGCC 3060  
Qy 2899 TGTGTATCTTCGAGAGACAGACCTGAAGAGCTGTGATCTCAGGACAAATGCTGTGGGT 2958  
Db 3061 TGTGTATCTTCGAGAGACAGACCTGAAGAGCTGTGATCTCAGGACAAATGCTGTGGGT 3120  
Qy 2959 GACGCTGGGCTGTGGCGTGTGCGAGAGAGCTGAAGAGAAAGAACATGTCTTGAACGAG 3018  
Db 3121 GACGCTGGGCTGTGGCGTGTGCGAGAGAGCTGAAGAGAAAGAACATGTCTTGAACGAG 3180  
Qy 3019 CTGCGATTGAAGGATGTGACTGATCTGATGCTGTGAGGAGCACTCTCTGGCCCTT 3078  
Db 3181 CTGCGATTGAAGGATGTGACTGATCTGATGCTGTGAGGAGCACTCTCTGGCCCTT 3240  
Qy 3079 TCTGCAACCGGCACTGACCACTGAACCTGTGTGAAGAAATTAAGTCCCAAGAG 3138  
Db 3241 TCTGCAACCGGCACTGACCACTGAACCTGTGTGAAGAAATTAAGTCCCAAGAG 3300  
Qy 3139 ATGATGAAGCTGTGTGGGCTTGTGCTGTCCACGCTTAACTTAAGATTAATGGGCTG 3198  
Db 3301 ATGATGAAGCTGTGTGGGCTTGTGCTGTCCACGCTTAACTTAAGATTAATGGGCTG 3360  
Qy 3199 TGAAGATGAGAGTACCTGTGCAAAATTAAGAAAGCTGTGAGAGAAAGTCACTCAAG 3258  
Db 3361 TGAAGATGAGAGTACCTGTGCAAAATTAAGAAAGCTGTGAGAGAAAGTCACTCAAG 3420  
Qy 3359 CCCGAGTCTGTAATTAAGCGGTAGTGGCATTTCTTTGATGAAGATGACCGAC 3312  
Db 3421 CCCGAGTCTGTAATTAAGCGGTAGTGGCATTTCTTTGATGAAGATGACCGGTAC 3474

RESULT 8  
us-10-216-645-1  
; Sequence 1, Application US/10216645

Publication No. US20030125282A1  
GENERAL INFORMATION:  
APPLICANT: WEISS, BERTRAM  
APPLICANT: LESSL, MONIKA  
APPLICANT: PETERS-KOTTIG, MICHAEL  
APPLICANT: BECKMANN, GEORG  
TITLE OF INVENTION: HUMAN MATER PROTEINS  
FILE REFERENCE: SCH-1910  
CURRENT APPLICATION NUMBER: US/10/216,645  
CURRENT FILING DATE: 2003-01-21  
PRIOR APPLICATION NUMBER: DE 101 39 874.3 41  
PRIOR FILING DATE: 2001-08-10  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 1  
LENGTH: 3926  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-216-645-1

Query Match 77.7%; Score 3133.6; DB 15; Length 3926;  
Best Local Similarity 95.2%; Pred. No. 0;  
Matches 3308; Conservative 0; Mismatches 4; Indels 162; Gaps 1;

QY 1 ATGGAAGGAGCAAAATCGCTCACCTTTCCAGCTACGGGCTGCATGATGCTCTATGAG 60  
DB 1 ATGGAAGGAGCAAAATCGCTCACCTTTCCAGCTACGGGCTGCATGATGCTCTATGAG 60  
QY 61 CTAGACAAGAGAATTTTCAAGATTTCAAGAAATTTACTAAGAGAATTTCTTCAAGATCG 120  
DB 61 CTAGACAAGAGAATTTTCAAGATTTCAAGAAATTTACTAAGAGAATTTCTTCAAGATCG 120  
QY 121 ACCACATGCTCTATTTCACAGTTTGAATTCAGAAATGCCACGTCGATGCTGCACTC 180  
DB 121 ACCACATGCTCTATTTCACAGTTTGAATTCAGAAATGCCACGTCGATGCTGCACTC 180  
QY 181 CTTCTGATGATATTATGAGACATCGCTGGCTGGGCTACATGATGATGATCTTTGAA 240  
DB 181 CTTCTGATGATATTATGAGACATCGCTGGCTGGGCTACATGATGATGATCTTTGAA 240  
QY 241 AACATGAACCTGCGAACCTCTCGAGAGAGGCAAGGATGATGATGATGATGATGATGAT 289  
DB 241 AACATGAACCTGCGAACCTCTCGAGAGAGGCAAGGATGATGATGATGATGATGATGAT 289  
QY 290 ----- 289  
DB 290 ----- 289  
QY 301 GATCCTGAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
DB 301 GATCCTGAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
QY 290 ----- 289  
DB 290 ----- 289  
QY 361 TATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
DB 361 TATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
QY 290 ----- 289  
DB 290 ----- 289  
QY 421 AAGTATGTTGAATTCATTTCTTTTTCAGAGAAATTTTCAAGATGATGATGATGATGATGAT 480  
DB 421 AAGTATGTTGAATTCATTTCTTTTTCAGAGAAATTTTCAAGATGATGATGATGATGATGAT 480  
QY 319 GCCACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 378  
DB 319 GCCACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 378  
QY 481 GCCACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 540  
DB 481 GCCACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 540  
QY 379 ACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 438  
DB 379 ACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 438  
QY 541 ACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 600  
DB 541 ACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 600  
QY 439 GTGATGACCAATTCGCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 498  
DB 439 GTGATGACCAATTCGCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 498  
QY 601 GTGATGACCAATTCGCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
DB 601 GTGATGACCAATTCGCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
QY 439 GACTGCGCGGAATGCAAAAGCTTGGCTGCTGCTTGTGATGATGATGATGATGATGATGATGAT 558  
DB 439 GACTGCGCGGAATGCAAAAGCTTGGCTGCTGCTTGTGATGATGATGATGATGATGATGATGAT 558  
QY 661 GACTGCGCGGAATGCAAAAGCTTGGCTGCTGCTTGTGATGATGATGATGATGATGATGATGAT 720  
DB 661 GACTGCGCGGAATGCAAAAGCTTGGCTGCTGCTTGTGATGATGATGATGATGATGATGATGAT 720  
QY 559 CCTGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618  
DB 559 CCTGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618

DB 721 CCTGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
QY 619 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 678  
DB 781 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
QY 679 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 738  
DB 841 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
QY 739 GAGTGGCCAGACCTCCAGAGCTCCGCTGAGAGATGATGATGATGATGATGATGATGATGATGAT 798  
DB 901 GAGTGGCCAGACCTCCAGAGCTCCGCTGAGAGATGATGATGATGATGATGATGATGATGATGAT 960  
QY 799 TTGATCATTTGACGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 858  
DB 961 TTGATCATTTGACGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
QY 859 AAAAGCTGGGCTGAGAACGACGCTCCGCTTCAACCTGATGATGATGATGATGATGATGATGAT 918  
DB 1021 AAAAGCTGGGCTGAGAACGACGCTCCGCTTCAACCTGATGATGATGATGATGATGATGATGAT 1080  
QY 919 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 978  
DB 1081 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
QY 979 TCAAGAGTGTGTCTCCCGCTTACCTGTTAGTGAAGAAATCTCCGAGGAAACAAAGATC 1038  
DB 1141 TCAAGAGTGTGTCTCCCGCTTACCTGTTAGTGAAGAAATCTCCGAGGAAACAAAGATC 1200  
QY 1039 CACTTGTCTCTTGAAGCGGAGATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098  
DB 1201 CACTTGTCTCTTGAAGCGGAGATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
QY 1099 ATGAACCAACCGTGAAGCTGCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1158  
DB 1261 ATGAACCAACCGTGAAGCTGCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
QY 1159 GTGGCCCTGACGCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1218  
DB 1321 GTGGCCCTGACGCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380  
QY 1219 ACAAGCTGACAGGCTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1278  
DB 1381 ACAAGCTGACAGGCTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
QY 1279 TGTCTCAATCTGAGAGAAAGATGCTGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1338  
DB 1441 TGTCTCAATCTGAGAGAAAGATGCTGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500  
QY 1339 GTGTGAATGAGAAAGATGCTGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1398  
DB 1501 GTGTGAATGAGAAAGATGCTGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560  
QY 1399 TGTGAGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1458  
DB 1561 TGTGAGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
QY 1459 TACTACACCTTCTTCAACCTGAGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1518  
DB 1621 TACTACACCTTCTTCAACCTGAGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680  
QY 1519 GAGGCGCTGGAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1578  
DB 1681 GAGGCGCTGGAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740  
QY 1579 ATGAGCTTTAAACAGGAGGCTTCAATATCAATGCTGCTTGTGATGATGATGATGATGATGAT 1638  
DB 1741 ATGAGCTTTAAACAGGAGGCTTCAATATCAATGCTGCTTGTGATGATGATGATGATGATGAT 1800  
QY 1639 GGCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1698  
DB 1698 GGCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1698



APPLICANT: Burgess, Catherine E.  
TITLE OF INVENTION: No. US004043382A1el Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-290C  
CURRENT APPLICATION NUMBER: US/10/092,900A  
CURRENT FILING DATE: 2002-03-07  
PRIOR APPLICATION NUMBER: USSN 60/274,322  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/283,675  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: USSN 60/338,092  
PRIOR FILING DATE: 2001-12-03  
PRIOR APPLICATION NUMBER: USSN 60/274,281  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/274,191  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/325,681  
PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: USSN 60/304,354  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: USSN 60/279,995  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: USSN 60/294,899  
PRIOR FILING DATE: 2001-05-31  
PRIOR APPLICATION NUMBER: USSN 60/287,424  
PRIOR FILING DATE: 2001-04-30  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 768  
SEQ ID NO 347  
LENGTH: 3226  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (69)..(3168)  
US-10-092-900A-347

Query Match 71.9%; Score 2900.6; DB 18; Length 3226;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2503; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 406 CATGAGGTGACACATGGGACTACAGAGTCACTGATGACCAATTCGCTGAGGAGG 465  
DB 249 CATGAGGTGACACATGGGACTACAGAGTCACTGATGACCAATTCGCTGAGGAGG 308  
QY 466 GATGACGTCTGATGTTTGAAGAACTGCTGCTGACGCGCGGAAATGCAACGTTGGCT 525  
DB 309 GATGACGTCTGATGTTTGAAGAACTGCTGCTGACGCGCGGAAATGCAACGTTGGCT 368  
QY 526 GGTGCTTTGATTCAGACCGGTGGGGCTTCGCGCTCCGACCGGTGTTCTGACGGGAAG 585  
DB 369 GGTGCTTTGATTCAGACCGGTGGGGCTTCGCGCTCCGACCGGTGTTCTGACGGGAAG 428  
QY 586 TCAGGAATTTGGGAATTCGGCTCTAGCCAGAAAGATCGTGTGCTGGGCGCAAGTGA 645  
DB 429 TCAGGAATTTGGGAATTCGGCTCTAGCCAGAAAGATCGTGTGCTGGGCGCAAGTGA 488  
QY 646 CTCTACCAAGGAATGTTCTCTACGTTCTTCTCTCCCGTTAGAGAGATGACGGGAAG 705  
DB 489 CTCTACCAAGGAATGTTCTCTACGTTCTTCTCTCCCGTTAGAGAGATGACGGGAAG 548  
QY 706 AAGGAGACAGTGTCAAGAGATTCATCTTCAGGAGTGGCCAGATCCCAAGCTCCGGTG 765  
DB 549 AAGGAGACAGTGTCAAGAGATTCATCTTCAGGAGTGGCCAGATCCCAAGCTCCGGTG 608  
QY 766 ACGGAGATCATGTCGCCAGCAGAAAGGCTGTGTTGATCATGACGTTTGGATGACCTG 825  
DB 609 ACGGAGATCATGTCGCCAGCAGAAAGGCTGTGTTGATCATGACGTTTGGATGACCTG 668  
QY 826 GGGCTGTCTCTCAACAATGACCAAAAGCTCTGCAAAAGCTGGGCTGAGAAAGAGCTTCG 885  
DB 669 GGGCTGTCTCTCAACAATGACCAAAAGCTCTGCAAAAGCTGGGCTGAGAAAGAGCTTCG 728  
QY 886 TTCACCTCATACGAGAGTCTGCTGAGGAAGTCTGCTCCCTGAGATCCTTCTGATCGTC 945

DB 729 TTCACCTCATACGAGAGTCTGCTGAGGAAGTCTGCTCCCTGAGATCCTTCTGATCGTC 788  
QY 946 ACCGTCAAGAGAGTGGGCAACAGAGAGTCAAGTACAGAGTGTCTTCCCGTTACTG 1005  
DB 789 ACCGTCAAGAGAGTGGGCAACAGAGAGTCAAGTACAGAGTGTCTTCCCGTTACTG 848  
QY 1006 TTAGTTAAGGAATTCGGGGGAACAAGAAATCCAGTGTCTCTGAGCGGGGATGGT 1065  
DB 849 TTAGTTAAGGAATTCGGGGGAACAAGAAATCCAGTGTCTCTGAGCGGGGATGGT 908  
QY 1066 GAGCATCAGAAAGACCAAGGTTGGTGCATCATGAAACAACCGTAGCTGCTCAGCAG 1125  
DB 909 GAGCATCAGAAAGACCAAGGTTGGTGCATCATGAAACAACCGTAGCTGCTCAGCAG 968  
QY 1126 TCCAGGTGCGCGCGGTGGGCTCTCTCATCTGCGTGCCTTCAGCTCAGAGCGTGGT 1185  
DB 969 TCCAGGTGCGCGCGGTGGGCTCTCTCATCTGCGTGCCTTCAGCTCAGAGCGTGGT 1028  
QY 1186 GGGGAGACGTCGCCCCCTTCAACCAACGCTCAACAGGCTTCAGCGCCGCTTTGTTT 1245  
DB 1029 GGGGAGACGTCGCCCCCTTCAACCAACGCTCAACAGGCTTCAGCGCCGCTTTGTTT 1088  
QY 1246 CATCAGCTCACCCCTCAGAGCGTGGTCCGCGCTGTCTCAATCTGAGGAAAGATTGTC 1305  
DB 1089 CATCAGCTCACCCCTCAGAGCGTGGTCCGCGCTGTCTCAATCTGAGGAAAGATTGTC 1148  
QY 1306 CTGAAGCGCTTCTGCGGTATGAGTGTGAGGAGTGTGAAATAGAACTCAGTGTGAT 1365  
DB 1149 CTGAAGCGCTTCTGCGGTATGAGTGTGAGGAGTGTGAAATAGAACTCAGTGTGAT 1208  
QY 1366 GGTGACGACCTTCATGTTTCAAGAGACTCGGGAGTGTGAGTCTCGTCTGTTTCAATG 1425  
DB 1209 GGTGACGACCTTCATGTTTCAAGAGACTCGGGAGTGTGAGTCTCGTCTGTTTCAATG 1268  
QY 1426 AACATCTTCTCCAGACAGCACTGTGAGGAGTCACTACCTTCTTCACTCAGTCTC 1485  
DB 1269 AACATCTTCTCCAGACAGCACTGTGAGGAGTCACTACCTTCTTCACTCAGTCTC 1328  
QY 1486 CAGGACTTCTGTGCGGCTTGTACTACGTTTGAAGGGCTCGAAATGAGCAGCTCTC 1545  
DB 1329 CAGGACTTCTGTGCGGCTTGTACTACGTTTGAAGGGCTCGAAATGAGCAGCTCTC 1388  
QY 1546 TGCCTCTGTACGTTGAGAAACAAGAGTTCATGAGCTTTAAACAGGCTTCAT 1605  
DB 1389 TGCCTCTGTACGTTGAGAAACAAGAGTTCATGAGCTTTAAACAGGCTTCAT 1448  
QY 1606 ATTCACCTGCTTGGATGAAGGTTTCTGTTGGCTCGTGAAGGAGTGAAGAGG 1665  
DB 1449 ATTCACCTGCTTGGATGAAGGTTTCTGTTGGCTCGTGAAGGAGTGAAGAGG 1508  
QY 1666 CCACCTGAGGCTCTGCTGGGCTGTCCGTTCCCTGCGGGTGAAGCAGAACTTCTGAC 1725  
DB 1509 CCACCTGAGGCTCTGCTGGGCTGTCCGTTCCCTGCGGGTGAAGCAGAACTTCTGAC 1568  
QY 1726 TGGGCTCTCTGTTGGGTCAAGAGCTTAATGCAACCCAGAGACACCTTGAAGCC 1785  
DB 1569 TGGGCTCTCTGTTGGGTCAAGAGCTTAATGCAACCCAGAGACACCTTGAAGCC 1628  
QY 1786 TTTCACCTGCTTTCGAGACTCAAGACCAAGAGTTGTTGCTTGGCACTTAAACGCTTC 1845  
DB 1629 TTTCACCTGCTTTCGAGACTCAAGACCAAGAGTTGTTGCTTGGCACTTAAACGCTTC 1688  
QY 1846 CAAGAAGTGTGGCTTCCGATTAAACAGAACTGAGCTTGAATGATCTTCTTGCCTC 1905  
DB 1689 CAAGAAGTGTGGCTTCCGATTAAACAGAACTGAGCTTGAATGATCTTCTTGCCTC 1748  
QY 1906 CAGCATGTCCGTATTTGGGAAATTCGGGTGATGTCAAAAGGATTTTCCCAAGAT 1965  
DB 1749 CAGCATGTCCGTATTTGGGAAATTCGGGTGATGTCAAAAGGATTTTCCCAAGAT 1808  
QY 1966 GAGTCCGCTGAGGAGATGCTGTGTGCTCTATGAGTGGGGATTAAGACCTCATGAG 2025

Db 1809 GAGTCCGCTGAGGATGTCCTGTGTGTCCTCTATGATGCGGATTAAGACCCCTCATAG 1868  
Qy 2026 GAGCAGTGGGAAGATTTCTGTCTCATGTTGGACCCACCCACCTGCGGACGTGAG 2085  
Db 1869 GAGCAGTGGGAAGATTTCTGTCTCATGTTGGACCCACCCACCTGCGGACGTGAG 1928  
Qy 2086 CTGGGACAGCAGCATCTTGACAGAGCGGCGCATGAAGACCTGTGTGCCAAGCTGAGCAT 2145  
Db 1929 CTGGGACAGCAGCATCTTGACAGAGCGGCGCATGAAGACCTGTGTGCCAAGCTGAGCAT 1988  
Qy 2146 CCCACCTGCAAGATTAACAGACCTGATGTTTGAAGAAATGACAGATTAACCCCTGTGTGAG 2205  
Db 1989 CCCACCTGCAAGATTAACAGACCTGATGTTTGAAGAAATGACAGATTAACCCCTGTGTGAG 2048  
Qy 2206 CACCTTGAGATGATGATGAGCCCAACCGTAACCTTAAGATCCCTCAACCTTGGAGGAC 2265  
Db 2049 CACCTTGAGATGATGATGAGCCCAACCGTAACCTTAAGATCCCTCAACCTTGGAGGAC 2108  
Qy 2266 CACCTGAAAGAGAGATGTAAGATGCGGTGTGAAGCTTTAAACACCCAAATGTTTG 2335  
Db 2109 CACCTGAAAGAGAGATGTAAGATGCGGTGTGAAGCTTTAAACACCCAAATGTTTG 2168  
Qy 2326 TTGAGATCTTTGAGGCTGATGTTGCTGTGATTTGACCCATGCTCTTACTGTAAGATCTCC 2385  
Db 2169 TTGAGATCTTTGAGGCTGATGTTGCTGTGATTTGACCCATGCTCTTACTGTAAGATCTCC 2228  
Qy 2386 CAAATCTTACGACCTCCCGACCGCTGAATCTCTGAGCCTGGCAGAAACAGGTGACA 2445  
Db 2229 CAAATCTTACGACCTCCCGACCGCTGAATCTCTGAGCCTGGCAGAAACAGGTGACA 2288  
Qy 2446 GACCAAGGAGTAAATGCTCTGATGATGCTTTGAGAGTCTCCAGTGGCCCTGAGAG 2505  
Db 2289 GACCAAGGAGTAAATGCTCTGATGATGCTTTGAGAGTCTCCAGTGGCCCTGAGAG 2348  
Qy 2506 CTGATCTGAGAGATGTTGGATCAACAGCAGGCTTGGCCAGATCTTGGCCCTGAGCCCTC 2565  
Db 2349 CTGATCTGAGAGATGTTGGATCAACAGCAGGCTTGGCCAGATCTTGGCCCTGAGCCCTC 2408  
Qy 2566 GTGAGCAACCGGAGCTTGACACACCTGTGCTTATCCAAACAGCCTGGGAGAGAGAGT 2625  
Db 2409 GTGAGCAACCGGAGCTTGACACACCTGTGCTTATCCAAACAGCCTGGGAGAGAGAGT 2468  
Qy 2626 GTAAATCTACTGTGTGATGATGAGAGCTTCCCACTGATGCTTGGAGAGGCTGATGCTG 2685  
Db 2469 GTAAATCTACTGTGTGATGATGAGAGCTTCCCACTGATGCTTGGAGAGGCTGATGCTG 2528  
Qy 2686 AATCAGTGCACCTGAGACCGGCTGTGTGTTTCTTGAACCTTGGCCCTTATGGGTAC 2745  
Db 2529 AATCAGTGCACCTGAGACCGGCTGTGTGTTTCTTGAACCTTGGCCCTTATGGGTAC 2588  
Qy 2746 TCATGCTGAGCAGCCTGAGCCTTATGATGAACCTTGTGAAGAACATGAGCGTGAAGCTT 2805  
Db 2589 TCATGCTGAGCAGCCTGAGCCTTATGATGAACCTTGTGAAGAACATGAGCGTGAAGCTT 2648  
Qy 2806 CTGTGCGAGGTCAATGAGAAACATCTTGTCACTCCAGACCTGAGTGTGTAAGTGT 2865  
Db 2649 CTGTGCGAGGTCAATGAGAAACATCTTGTCACTCCAGACCTGAGTGTGTAAGTGT 2708  
Qy 2866 CATCTACCGCGCGGTGTGTGAGAGTGTGTCTGTGTGATCTGAGAGAGACAGACCTG 2925  
Db 2709 CATCTACCGCGCGGTGTGTGAGAGTGTGTCTGTGTGATCTGAGAGAGACAGACCTG 2768  
Qy 2926 AAGAGCCTGATCTCAAGGACATGCTTGGGTGAAGGTTGGGTTGCTGCGTGTGAG 2985  
Db 2769 AAGAGCCTGATCTCAAGGACATGCTTGGGTGAAGGTTGGGTTGCTGCGTGTGAG 2828  
Qy 2986 GAGCTGAAGCAAAAGACAGTGTCTGACAGACTCGGGTGAAGGCAATGTGACTGACT 3045  
Db 2829 GAGCTGAAGCAAAAGACAGTGTCTGACAGACTCGGGTGAAGGCAATGTGACTGACT 2888  
Qy 3046 TCTGATGCTGTGAGGACCTCTCTTGGCCCTTTCTGCAACCGGACATCTGACAGTCTA 3105  
Db 2889 TCTGATGCTGTGAGGACCTCTCTTGGCCCTTTCTGCAACCGGACATCTGACAGTCTA 2948

Qy 3106 AACCTGGGAGATTAATCTTCACTGCCAAAGATGATGAAGCTGTGTGGCCCTTTGCC 3165  
Db 2949 AACCTGGGAGATTAATCTTCACTGCCAAAGATGATGAAGCTGTGTGGCCCTTTGCC 3008  
Qy 3166 TGTCCACGCTTAATCTTACAGATTAATGGGCTGTGGAATGAGCATCTCTGTGCAATA 3225  
Db 3009 TGTCCACGCTTAATCTTACAGATTAATGGGCTGTGGAATGAGCATCTCTGTGCAATA 3068  
Qy 3226 AGAAGCTGCTGAGAGAGTGAAGCTACTCAAGCCCGGAGTGTATTTGACGAGTGTGG 3285  
Db 3069 AGAAGCTGCTGAGAGAGTGAAGCTACTCAAGCCCGGAGTGTATTTGACGAGTGTGG 3128  
Qy 3286 CATCTTTTATGATGAAGATGACCGACAC 3312  
Db 3129 CATCTTTTATGATGAAGATGACCGGTAC 3155

RESULT 10  
US-10-399-443-5  
; Sequence 5, Application US/10399443  
; Publication No. US2004002869A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as Represented by the  
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
; APPLICANT: Nelson, Lawrence M.  
; APPLICANT: Nelson, Zhi-Bin  
; APPLICANT: Tong, Zhi-Bin  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Zhi-Bin, Tong  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64785  
; CURRENT APPLICATION NUMBER: US/10/399,443  
; CURRENT FILING DATE: 2003-04-16  
; PRIOR APPLICATION NUMBER: 60/241,510  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; PRIOR FILING DATE: 2001-04-04  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 3447  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-399-443-5

Query Match 31.7%; Score 1278; DB 17; Length 3447;  
Best Local Similarity 67.2%; Pred. No. 0;  
Matches 1898; Conservative 0; Mismatches 908; Indels 18; Gaps 6;

Qy 498 TGACTGGCGGGAATGCAAAAGTTGGCTGTTGATTGATTCAGACCGTGGGCTTCG 557  
Db 515 TGAAGCCGCAAGATGAATTAATTTGTCTGATGCTTTTAACATACAGAAAACCTTCCA 574  
Qy 558 GCTTCGACCGGTGTTCTGACCGGAAGTCAAGAAATGCGACTTAAGCAGAAG 617  
Db 575 GCTTCACACCATTAATCTTACATGAGAACAGAGAGTGGAGTCAAGCTTTGGCAGAAG 634  
Qy 618 GATGCTGTGTGTGGCGGCAAGTGAATCTTACAGGAATGTCTTCTACGCTTCTT 677  
Db 635 TATGTTCTTGGCTGGGCAAGGATTAATCTTCCAAAATG--TCTTTTGTCTATCTT 691  
Qy 678 CCTCCCGTTAGAGAGATGACGCGGAAGAGAGAGAGAGTGTCAAGAGTTCACTCCAG 737  
Db 692 CTCTTCTGTTAGAAATTAAGTGAAGAGAGAGAGAGAGTGTGACACAGCTGATTTGCTTA 751  
Qy 738 GAGATGCAAGACTCCAGGCTCCGATGACGAGATCATGTCCGACAGAAAGCTGTT 797  
Db 752 GAGATGCAAGACTCCGAGATCTAATGACAAAGATCATGTCCGACAGAAAGACTCTT 811  
Qy 798 GTTATCATTTGACGTTTCCATGACCTGGGCTGTCTCTC--AAATGACAAAGCT 854  
Db 812 GTTGTATGATGAGCTTGTGATGATGACTGTGTCTCCACATGATGATGACACT 871

QY	855	CTGCAAAAGATCGGGGCTGAGAAAGACACCTCCCTTCAACCTCCATACCGATGCTCTGAGGAA	914
Db	872	ATCCAGAGACTGGAAAGATGAAACAGCCCATATACATCTCGATGTACAGCTCTTGAGGAA	931
QY	915	GGTCCCTGCTCCCTGAGTCCTTCTGATCGTCAACCGTCAGAGACGTGGGACAGAGAGCT	974
Db	932	GGCCTCTTACTCACTGCTCTTCTCATATATACACACGAAACACAGGCTTAGAAAACT	991
QY	975	CAAGTCAGAGGTCCGTGTCTCCCGTTACTGTTAGTAGGAAATCTCCGGGAAACAAG	103
Db	992	CAAGTCAAATGGTTGTGTCCCTCCCTATATACCTGGTTGAAGACCTGTGCAATCAAGAG	105
QY	1035	AATCACTTGTCTCTTGAGCGGGGATTTGGTAGAGATCGAAGAGACAAAGGTTGGTGTG	109
Db	1052	ATCTCACTGGTCTCTCGAAGAACTCTCCATATAGTCTGATAGAAATACAGTCTTCATTCC	111
QY	1095	GATCAATGAACAACCGTAGCTGCTCGACCAATGTCAGAGTGCCCGCTGGGCTCTTCAT	115
Db	1112	TCTGTATGAATAATCAACAGCTGTTTGAACCAATGCAGGCGCCCTCTGTGTGCTCCCTGG	117
QY	1155	CTGCGTGGCCCTGCGACTGTCAGAGAGTGGTGGGGAGAGCGTCCGCCCTTCAACAAAC	121
Db	1172	CTGTAGGCTCTACAGCTACAAAGAAACTGGGAAAAGATGACACCTTACCTGTCCAGAC	123
QY	1215	GCTCAAGGCTGCAACGCGCTTTTGTGTTTATAGCTCAACCCCTCAGGCGTGGTCCG	127
Db	1232	TCTCAACGGTTTGTATGTCAGCTGTGTGTTTACACAGCTCACTTTGAAAAGGCTTTCCA	129
QY	1275	GCGCTGTCTCAATCTGAGAGAAAGAGTTGTCTGAAGCGCTTCTGCGTATGGCTGTGA	133
Db	1292	GAGCCCTCTCACTCAGGAAGAACAGATTACTTAAGTGGTGTGTGTGATGATGTGACGTGA	135
QY	1335	GGAGATGTGGAATAGAAATCAGTGTGTTGATGATGTCGACCTCATGGTTCAAGACCTGG	139
Db	1352	AGGATGTGGAACATGAGGTCCGTGTTCTATATATAGACTCGAAGAACTATAGCTTAA	141
QY	1395	GGAGTCTGAGCTCCGTGCTCTGTGTTTCAATGAACAATCCTTCTCCAGACAGCC--ACTG	145
Db	1412	GGAGCTGAGATCTTGGCCCTCTTTCATGAACAATCCTTCTCCAGTGGTGGCAACAAG	147
QY	1452	TGAGGATGATACACACTTCTTCCACTCAGTCCAGATCTTCTGTGCGGCTTGTACTA	151
Db	1472	TGAGGATGTTATGTTTCTTCCCACTCAGCTTGAGAGATTTCTTGTGCTCTTATATTA	153
QY	1512	CGTGTTAGAGGGCTGGAATGAGCCAGCTCTGCGCTCTGACGTTGAGAAACAAA	157
Db	1532	TGTTTATGAAGGCTGGA--GGAATGAATATAGCATTTTGTCTCATTGAAACCAAG	158
QY	1572	GAGGTCCATGAGCTTAAACAGGCAAGCTTCCATATCCAATCGCTTTGATGAAGCTTT	163
Db	1589	GAGCATCATGAGGGTGAAGAGAACATGACGACACTGCGCTC--CTCGGATATMACGTTT	164
QY	1632	CTGTGTTTGGCCCTCGTAGAGGAAAGATGAAGGAGCGCACTGAGAGTCTGTGGGCTGTCC	169
Db	1646	CTTATTTGGCTCATGAACAGAGATATCTTGAAGACTGTGAAGGTTCTGTGTTGAATATCC	170
QY	1692	CGTTCCTCGGGGGTGAAGCAGAACTTCTGCACTGGGTCTCTCTGTTGGGTCAAGACGC	175
Db	1706	CGTAGTTCCAACTGTTTGAAGAGAACTCCAACATCGGGTCTCTGATAGCTCAAGAGGT	176
QY	1752	TAAATGCCAACCCAGAGAGAACCTTGAAGCGCTTCCACTGTCTTTTGAAGCTCAAGA	181
Db	1766	CAATGGAACAGCCCAATGGAACACCTGTGATGCTTCAATGTCTAATTTGAATCTCAGGA	182
QY	1812	CAAAAGATTTGTTCCCTTGGCAATTAAACAGCTTCAAGAAAGTGGCTCCGATTAACCA	187
Db	1826	TGAAGAGTTTGTGTGGCGGGGCTCTCAAAAGCTTCCAAGAAAGTGTGGCTGTGATTAACA	188
QY	1872	GAACTGGAATGATAGCATTTCTCTTGTGCTCAAGCATGTGCTCGATTTGGGAAAT	193
Db	1886	GAAAGTGAATGGAAGTCTCTCTTCACTGTCTCAAGCATGTGCAAACTTGAAAGCAAT	194

OY	1932	TCGGGTGGAATGCAAAAGGATCTTCCAGAGAGATGAGTCCGCTGAGGCAATGCTCGTGGT	1991
Db	1946	CCGGGTGATATACAGAGACCTCTCTCTCGGTAGATATACTCTGAGCTGTGCCCTGTGT	2005
OY	1992	CCCTCTATGATGATGGGGA---TAAGACCCCTCATTTGAGGAGCAGTGGAGATTTCTGCTC	2048
Db	2006	TACTGTCCAGAGACACATGTTAGGCCCTCTCTCATGAGATGTGTGGGAAACTTCTGCTC	2065
OY	2049	CATGCTTGGACCCACCCACCTGCGGACCTGGACCTGGGCGAGCAGCATCTTGACAGA	2108
Db	2066	TGTGCTTGGACGAGCTCCGGAACTTTGAAGAGACTGGACTTGTGGCGCAGACATCTTGACGA	2125
OY	2109	GGGGGCCATGAAAGACCTGTGTGTGCCAAGCTGAGGCACTCCACCTTGCAAGATATACAGACCT	2166
Db	2126	ACGGGCCCATGAAAGATACGTGTCCCTGAGCTCGGAATGATGCTTGCAAAATACAGAAACT	2185
OY	2169	GATGTTTGAATGACAGATTAACCCCTGGTGTGACAGCACTCTGAGAAATATGTCATGGC	2222
Db	2186	GACGTTTAAAGTGCAGAGGTAGTGTGTGTGGCCTGAACATCTCTGGAAGCTTCCTTTTAA	2245
OY	2229	CAACCGTAACTTAAGATCCCTCAACTTGGAGGACACCACTGGAAGAAAGATGTAA	2288
Db	2246	CAATCAAAACTTAAGATACCTCAATCTTAGGAAACATCCATGAAGATGATGACATGA	2305
OY	2289	GATGGCGTGTGAACCTTTAAACCCCAAAATGTTTGTGGAATCTTTGAGCGTGAATG	2348
Db	2306	GTTAGCCCTGCACAAACGCTGAAACATCCAAATGTCTCGTGAAGACTCTGAAGTTTGATTC	2365
OY	2349	CTGAGATATGACCATGCTGTGTAACCTGAAGATCTCCCAATCTCTACGACTCCCCAG	2408
Db	2366	CTGTAGTTAAACATCATTTGTTATGAGATGATCTCCAGCTTCTTATTTCAACACAG	2425
OY	2409	CCTGAAATCTCTGAGCCTGCGCAGGAAAACAAGTGAACAGACGAGGAATGACCTCTGAG	2468
Db	2426	GCTAAAGTGTCTCAGCCTGCGCAAAAATAGCTGGAGTAAAAAGCATGATATCCCTTGG	2485
OY	2469	TGATGCTTGAAGACTCTCCAGTGGCCCTGCAAGAGCTGATCTGAGAGCTGTGGCAT	2528
Db	2486	GAAATGCCCTTGAATGCTCAATGTGTCTAATGCAAAAGTTGATCTGCAACATGTGGCT	2545
OY	2529	CACAGCCACGGGTTGCCAGAGTCTTGCGCTGACCCCTGTGAGCAACCGAGACTTGACA	2588
Db	2546	CACACTGCGCAGCTGCACCTTCTGTGCTCAGCCCTTTTCAGCAACCAAGACTTGACA	2605
OY	2589	CCTGAGCCTATCCAAACAACAGCCTGCGGGAACGAAGTGTAAATCTAATGTGTGATCCAT	2648
Db	2606	CTGTGCTGTCTCAACAACAAGCCTTGGGGACTGAAGAGTGCACAGCTGTGTACGTTCT	2665
OY	2649	GAGGCTTCCCACTGTAGTGTGACAGAGGCTGATCTGATCATGTCACCTGACACGGC	2708
Db	2666	GAGGAATTCAGAAATGTGTCTCCAGCGGCTGATACTGATATCTGCAACATTTGATGA	2725
OY	2709	TGCGCTGTGTTTTCTTGCACTTGGCGCTTATGTGGTAACTCATGTGCTGACGACCTGAGCT	2768
Db	2726	TGCTTATGTGCTCTCTGCAATGAGACTTGCACAAACAAGAGTGAACCACTGAGGCT	2785
OY	2769	TAGCATGAACCTGTGGAAGACATGTGGCGTGAACCTTGTGCGAGGCTCATGAGAGAAC	2828
Db	2786	GACCATTAACCCGTAAGGAGTGTGTCAATGAACTCTGTGTGAAGCTTTAAAGAAAC	2845
OY	2829	ATCTTGTATCTCCAGAGACTGTGAGTTGTAAATGTATCTCAACCGCGCGTGTGTGA	2888
Db	2846	TACTTGTATCTTCAAGAACTGGAACATAGTGAATGCAACTCAACAGAACTGTGGGA	2905
OY	2889	GAGTCTGTCTGTGTGATCTTGAGAGAGACACCTGAAGAGCTGTGATCTACGACAA	2948
Db	2906	GGACTGTGCGCTGTATGTATCAACAAACCAAGACCTTAAAAAGTTTGATCTTGTGTAACA	2965
OY	2949	TGCGCTGGTGAACGCTGTGGGTTGTGTGAGCTGTGGAAGGGACTGAAGCAAAAGACAGTGT	3008
Db	2966	CGCCCTGTGGTGAACAAAGAGTCAATACCTGTGTGTGAAGGGACTGAAGCAAAAGTACAGCTC	3025
OY	3009	TCGACACAGACTCGGGTTTGAAGGCAATGTGACTGATCTGATTTGCTGTGTGAGGCACTCTC	3068



[illegible]

## RESULT 11

US-10-677-943--5  
 ? Sequence 5, Application US/10677943  
 ? Publication No. US2004007237A1  
 ? GENERAL INFORMATION:  
 ? APPLICANT: The Government of the United States of America as  
 ? APPLICANT: represented by the Secretary of the Department of Health and  
 ? APPLICANT: Human Services  
 ? APPLICANT: Nelson, Lawrence  
 ? APPLICANT: Tong, Zhi-Bin  
 ? TITLE OF INVENTION: Human Gene Critical to Fertility  
 ? FILE REFERENCE: 4239-64790  
 ? CURRENT APPLICATION NUMBER: US/10/677,943  
 ? CURRENT FILING DATE: 2003-10-01  
 ? PRIOR APPLICATION NUMBER: 60/241,510  
 ? PRIOR FILING DATE: 2000-10-18  
 ? PRIOR APPLICATION NUMBER: PCT/US02/09776  
 ? PRIOR FILING DATE: 2002-03-29  
 ? PRIOR APPLICATION NUMBER: PCT/US01/10981  
 ? PRIOR FILING DATE: 2001-04-04  
 ? NUMBER OF SEQ ID NOS: 42  
 ? SOFTWARE: PatentIn version 3.2  
 ? SEQ ID NO 5  
 ? LENGTH: 3447  
 ? TYPE: DNA  
 ? ORGANISM: Mus musculus  
 ? US-10-677-943--5

Query Match 31.7%; Score 1278; DB 18; Length 3447;

Matches 1898; Conservative 0; Mismatch

Matches 1898; Conservative 0; Mismatches 908; Indels 18; Gaps 6;

	498	515	558	575	618	635	678	692	738
OY	TGACTGGCCGGAATGCAAA	CGTTGGCTGGTCTTTGATTC	CAGACCGAGTGGGCGTTCCG						
Db	TGACGAGCCCAAGATGAAAT	TATTCTGTGATCTTTTAAC	CATACAGAAAACTTTCA						
OY	558	GCCTTGCACGGTGTTC	TGCACGGAAAGTCAGGAAT	TGGAAATGGCTCTT	TGCCAAG	617			
Db	575	GCCTTGCACCAATTATCT	TCAATGGAAGACAGAGAT	TGGGAAGTAGCTTTGG	CCAGAAG	634			
OY	618	GATCGTGTGCTGCTGGG	CGCAAGTGTGACTTAC	CAGGAAATGTCTCTTA	CGTCTTCTT	677			
Db	635	TATTGTCTTGGCTGGG	CACAGGGTAACTCTT	CCAAAAAAG---	TCTTTGTATCTT	691			
OY	678	CCTCCCCCTTGAAGATG	ACCGGGAAGAAGAGCAG	TGTGCACAGAGTTCA	TCTCCAG	737			
Db	692	CTTCTCTTTAGAAATTA	AGTTGACAGAAAGCAG	ATTGGGACACGCTGAT	TTCATA	751			
OY	738	GGAAGTGACAGCTTCC	CAGGCTCCGAGTAC	CGGAGATCTGTCTCC	GACCAAGAAAGGCTTT	797			

Db	752	GGAGGTTCAGACTTCCTGGGATCTAGTGAACAAGATCATGTCCCAACGAGAAAGCTCTT	811
Oy	798	GTTCATCATGTAACGGTTTTCATGACCTGGGCTGTGTCTC---AACAATGACACAAAGCT	854
Db	812	GTTTGTATAGATATGGCTTGATGATGATAGACCTGTGTCTCCAAACATGATGATAGACCT	871
Oy	855	CTGCAGAAAGACTGGGCTGAAGACAGCCTCCGTTCAACCTCATACGACGTCTGTGAGGAA	914
Db	872	ATCCAGAGACTGGGAAGGATGAAACGCCCAATATCATCTGATGTAAGCCTCTCGAAGAA	931
Oy	915	GGTCCTGTCCCTGAGTCTTCCATGATGTCACCGTCAGAGAGCGTGGGACAGAGAAAGCT	974
Db	932	GGCTCTCTTACTTCAAGTCTTTCTCATCATTAATCCACGAAACACAGGCTTGAAGAACT	991
Oy	975	CAAGTCAGAGTCTGTCTCTCCCGTTACCTGTAAGTAGAATCTCCGGGAAACAAG	1034
Db	992	CAATTCATATGGTTGTGTCTCCCTCTATATATCTGTTGAAGGACTGTCTGATCAAGAG	1051
Oy	1035	AATCCACTTGTCTTTAGCGCGGGATTGGTAGCATCAAGACACAGGGTTGCGTC	1094
Db	1052	ATCTCAGCTGGTCTCTCAGAAACATCTCCAAATGAGTCTGATAGAAATACAAGCTTCCATTC	1111
Oy	1095	GATCATGAACAACCGTAGCTGTCTCCACCATGTGCAGTGTGCGCGCGGTGGCTCTCAT	1154
Db	1112	TCTATAGAAATATACCAAGCTGTTTGACCAATGCAAGGCCCTCTGTGTCTCTGTGT	1171
Oy	1155	CTCGGTGGCCCTGAGCTGACAGGACGTGTGTGGGGAGAGCGTCGCCCTTCAACCAAC	1214
Db	1172	CTGTAGAGGCTCTAAGCTTACAGAAAGAACTGGGAAAGATGACCTTACCTGTCCAGAC	1231
Oy	1215	GCTCAACAGGCTGCAGCGCCGCTTTGTGTTCATCAGCTCAACCTCGAGGCGTGTCCG	1274
Db	1232	TCTCAACGGTTTGTATGCCAGTGTGTGTTCACCACTCACTTGAAAGGCTTCCCA	1291
Oy	1275	GCGCTGTCTCAATCTGAGAGAAAGATGTTCGTAAGCGCTTCTGCGGTATGGCTGTGA	1334
Db	1292	GAGGCTCTCAGTAGAGAAACAGATTACTTATGTGGTTGTGCAATGATGGAGCTGA	1351
Oy	1335	GGGAGTGTGAATATGAAAGTCAAGTGTGTGATGTGAGACACTCAATGTTCAAGACTCG	1394
Db	1352	AGGAGTGTGACCATGAGTGTGTGTCTATGATGATGACTGAAGAACTATAGCTTAA	1411
Oy	1395	GGAGTGTGAGTCCGTGCTCTGTTCATCATGAACATCTTCTCTCCAGACAGCC---ACTG	1451
Db	1412	GGAGTCTGAGATCTTGGCCCTCTTTCACATGAACATCTTCTCCAGGTGTGGCCACAACAG	1471
Oy	1452	TGAGAGATATCATCACTTCTTCCACTCACTGTCACAGACTTCTGTGCGCCCTGTGACTA	1511
Db	1472	TGACCAAGTGTATGTTTCTTCCCACTCAGCTGACAGAAATTTCTTGTCTCTCTTATATTA	1531
Oy	1512	CGTGTAGAGGGCTGGAATAATCGACGAGCTCTCTGCGCTCTGTGACCTTGAGAGACAA	1571
Db	1532	TGTTTGAAGAGGCTGGA---GGAATGAAATCAGCAATTTTGTCTCATTTGAAAACAAAG	1588
Oy	1572	GAGGTTCATGAGCTTAAACAGGAGGCTTTCATATCCACTCGTTTGTGATGAAGCTTT	1631
Db	1589	GAGCATCATGAGGTGGAAGAAACTGACGCACTGCGCTC---CTCGGAAATGAAGCTTT	1645
Oy	1632	CTTGTGTGGCTCTGTGAGCGAAGCTGTAAGAGGCCACTGGAAGTCTCTGCTGGGCTGTCC	1691
Db	1646	CTTATTTTGGCTCTCAATGAACAAGATATCTTGAAACACTGGAAGGTTGTGTTGAATATCC	1705
Oy	1692	CGTTCCTCCGAGGAGTGAAGAGAACTTTCGACCTGGAGTCTCTGTGTGGGTGACAGGCC	1751
Db	1706	CGTATATTCACATCTTTAGAGAGAGCTTCAACACTGGAGTCTCTGTGATAGCTCAGCAGGT	1765
Oy	1752	TAAATGACACACCCGAGAGACACCTTGAACGCTTCACTGTCTTTTTCAGACTCAAGA	1811
Db	1766	CAATGGGACACAGCCCAATGAGACACCTGGATGCTTATTTGTCTATTTAGAGTTCAGGA	1825
Oy	1812	CAAAAGTTTGTTCCTTGGCATTTAAACAGCTTCCAAAGAAATGTGGCTTCCGATTTAAACA	1871

Db 1826 TGAAGATTGTTGGCGGGCTCTCAACGCTTCAAGAGTGTGCTGATTACCA 1885  
 Qy 1872 GAACCTGGAATTGATGATCTTCTTGCCTCCGACCTGCTCCCTATTGGGAAAT 1931  
 Db 1886 GAAAGATGGAATTGAAAGTCTCTTCTTCACTGTCTCAAGACCTGCAACCTTGAAGCAAT 1945  
 Qy 1932 TCGGGTGAATGCAAAAGGATCTTCCCAAGAGATGATCGGCTGAGGACATGCTGTGGT 1991  
 Db 1946 CCGGGTGAATGCAAGACCTCTCTCGGATGATTAATCTCTGAGACTGTGCTGTGT 2005  
 Qy 1992 CCTCTATGAGATGCGGGA--TAAACCTCATTTAGAGAGACAGTGGAAATTTCTGCTC 2048  
 Db 2006 TACTGTCCAGAGACACAATGTATAGCCCTCTCTCATGAGATGGTGGGAAACTTCTGCTC 2065  
 Qy 2049 CATGCTGGCAACCCACACCTGCGGAGCTGACCTGGGAGAGACATCTTGACAA 2108  
 Db 2066 TGTGCTGGCAGCTCCGGAATCTTGAAGAGCTGAGCTTGGGGAGACAGATCTTGATCA 2125  
 Qy 2109 GCGGGCATGAAAGACCTGTGTGCCAAGCTGAGAGCTCCACCTGCAAGATACAGACCT 2168  
 Db 2126 ACGGGCATGAAAGATATCTGTGCTCGAGCTCGGGAATCACTCTGCAAGATACAGAGCT 2185  
 Qy 2169 GATGTTAGAAATGCAAGATTAACCTGTGTGACGACCTCTGAGAAATGCTCATGCG 2228  
 Db 2186 GAGTTTAAAGAGAGAGAGTAGTGTCTGGCTGAAACATCTCTGGAAGCTCTTTTAG 2245  
 Qy 2229 CAACCTAACTTAAGATCCCTCAACTTGGGAGGACCCACCTGAGAAAGAGATTTAG 2288  
 Db 2246 CAATCAAACTTAAGATCTTCACTTACCTAGGAAACATCCCATGAAAGATGACATGAA 2305  
 Qy 2289 GATGGGATGAAAGCTTAAACACCCAAATGTTTGGATGTTGAGGCTTGGAGCTGATG 2348  
 Db 2306 GTTAGCTGCAAGAGCTGAAACATCAAGTCTCGTAGAGACTCTGAGGTTGATTC 2365  
 Qy 2349 CTGTGATTTGACCCATGCTCTTAACTGTAAGATCTTCCAAATCTTACGACCTCCCGAG 2408  
 Db 2366 CTGTGATTTAACCATGATGTTTATGATGATCTCAGGCTTCTTATTTCAACGACGAG 2425  
 Qy 2409 CTTGAATTTCTGAGCTGCGGAGAAACAGGTGACAGACAGGAGTAAATGCTCTGAG 2468  
 Db 2426 GCTAAAGTGTCTCAGCTGCGGCAAAATAGAGTGAATAAAGATATATCCCTTGG 2485  
 Qy 2469 TGATGCTTGAAGTGTCCAGTGGCGCTGCAAGAGTGAATCTGAGAGACTGTGGCAT 2528  
 Db 2486 GAATGCTTGAAGTGTCAATGTGTCTACTGCAAAATGTGATCTGGAACACTGTGGCT 2545  
 Qy 2529 CACAGCAAGGATGCGCAGAGTGTGCTCAGGCTCTGACAGAACCGAGCTTGAACA 2588  
 Db 2546 CACACTGTGACGTGCACTTCTGTGTCTCAGGCTTTTACAGCAACGAACTTGACACA 2605  
 Qy 2589 CCTGTGCTATCCAAACAGCCTGGGGAACGAAAGTGTAAATCTAATGTGTGATCCAT 2648  
 Db 2606 CCTGTGCTGTCAACAGCCTGGGGAACGAAAGTGTGCAACAGCTGTGTGATGCT 2665  
 Qy 2649 GAGGCTTCCCACTGTGTGCAAGAGTGTGATGTGATCTGATCTGACACTTGACAGCGC 2708  
 Db 2666 GAGGAATTCAGAAATGTCTCTCAGGCTGTGATCACTGAAATCACTGAAATGTATGATA 2725  
 Qy 2709 TGGCTGTGTTTCTTCTGCACTTGGCTTATGGTAACTCATGAGCTGACACCTGAGCT 2768  
 Db 2726 TGTCTTATGCTTCTCTGCAATGAGACTTTCGAAACAGCAAGAGTGAACCTGAGCT 2785  
 Qy 2769 TAGCATGAACCTGTGGAAGACATGGCTGAGCTTCTGTGCGAGGTCATGAGAAAC 2828  
 Db 2786 GACCATGAACCCCGTGGGAGTGTGCAATGAAAGCTATGTGAAAGCTTTTAAAGAAC 2845  
 Qy 2829 ATCTTGTCTCTCCAGGACCTGAGAGTGTGTAAGTGTCTTCAACGCGCGGCTGTGTGA 2888  
 Db 2846 TACTTGTATCTTCAAGAACTGAACTGAGCTGCAACTTCAACAGAACTGTGCGCA 2905  
 Qy 2889 GAGTCTGTCTGTGTGATCTGAGAGAGACACCTGAAAGAGCTGATCTCAAGGACAA 2948  
 Db 2906 GGAACCTGTGTATGATCAACAAACAGCACTTAAAGTTTGGATCTTGTGTACAA 2965

Qy 2949 TGCCTGGGTGACGGTGGGGTGTCTGCGCTGTGCGAGGACCTGAAGCAAAAGACATGT 3008  
 Db 2966 CGCCTTGGGTGCAAAAGAGATATTAACCTGTGTGAGGAGCTGAAGCAAAAGTACAGCTC 3025  
 Qy 3009 TCTGAGAGACTCGGGTGTGAAGCATGTGACATCTTCTGATTTGCTGTGAGGACATCTC 3068  
 Db 3026 CTTGAGAGACTTGGGTTGGGGCATGTATGATGATCTTCAATTTGCTGTGAGGACATCTC 3085  
 Qy 3069 CTGGGCTTCTCTGCAACCGGACATGTGACATCTTAAACCTGTGTGAGGACATCTTCA 3128  
 Db 3086 ATTTGCAATCTTTCACACCTTCACTGAAACAGCTTAAACCTGTGAGGAAATGATCTTCA 3145  
 Qy 3129 TCCCAAGGAATGATGAAGCTGTGTGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 3188  
 Db 3146 TACATCGGGAGATGTGAAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3205  
 Qy 3189 AATTGGGCTGTGGAATGTGCACTGATACCTGTGCAATTAAGAAAGCTGTGAGAAAGTGA 3248  
 Db 3206 AATTGGCTGTGGAAGCAAGAGTATATGCTGCAAGTGAAGAAAGCAAGCTGTGAGAAAGTGA 3265  
 Qy 3249 GCTACTCAAGCCCGAGTGTGATTAATGACGTAATGTGCAATTTTGTATGAAGATGACCG 3308  
 Db 3266 GTTGTCAAGCCCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 3325  
 Qy 3309 ACAC 3312  
 Db 3326 AAC 3329

RESULT 12  
 US-10-399-443-1  
 ; Sequence 1, Application US/10399443  
 ; Publication No. US20040028659A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The Government of the United States of America, as Represented by the  
 ; APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
 ; APPLICANT: Health  
 ; APPLICANT: Nelson, Lawrence M.  
 ; APPLICANT: Tong, Zhi-Bin  
 ; APPLICANT: Nelson, Lawrence  
 ; APPLICANT: Zhi-Bin, Tong  
 ; TITLE OF INVENTION: Human Gene Critical to Fertility  
 ; FILE REFERENCE: 4239-64785  
 ; CURRENT APPLICATION NUMBER: US/10/399,443  
 ; PRIOR FILING DATE: 2003-04-16  
 ; PRIOR APPLICATION NUMBER: 60/241,510  
 ; PRIOR FILING DATE: 2000-10-18  
 ; PRIOR APPLICATION NUMBER: PCT/US01/10981  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 1  
 ; LENGTH: 1157  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-399-443-1

Query Match 28.6%; Score 1152.8; DB 17; Length 1157;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1154; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 754 CAGGCTCCGGTGAACGAGATCATGTCCGACAGAAAGCTGTGTTTCATCATTTGACGCT 813  
 Db 1 CAAAGCTCCGGTGAACGAGATCATGTCCGACAGAAAGCTGTGTTTCATCATTTGACGCT 60  
 Qy 814 TTGCATGACCTGGGCTGTGTCTCTCAACATGACACAAAGCTCTGCAAAAGACTGGGCTGAG 873  
 Db 61 TTGCATGACCTGGGCTGTGTCTCTCAACATGACACAAAGCTCTGCAAAAGACTGGGCTGAG 120  
 Qy 874 AAGGAGCTCCGTTCAACCTCATAGAGAGTGTGCTGAGGAAAGTCTGCTCCCTGAGTCC 933  
 Db 121 AAGGAGCTCCGTTCAACCTCATAGAGAGTGTGCTGAGGAAAGTCTGCTCCCTGAGTCC 180

934 TTCTGATGTCACCGTCAGAGACGTGGGACAGAGAAGCTCAAGTCAGAGGTGTGCT 993  
Db TTCTGATGTCACCGTCAGAGACGTGGGACAGAGAAGCTCAAGTCAGAGGTGTGCT 240  
994 CCCCGTACCTGTTAGTTAGAGAAATCTCCGGGAAACAAAGATCACTTGTCTTGG 1053  
Db CCCCGTACCTGTTAGTTAGAGAAATCTCCGGGAAACAAAGATCACTTGTCTTGG 300  
994 CCCCGTACCTGTTAGTTAGAGAAATCTCCGGGAAACAAAGATCACTTGTCTTGG 1053  
Db CCCCGTACCTGTTAGTTAGAGAAATCTCCGGGAAACAAAGATCACTTGTCTTGG 300  
1054 CGCGGATTTGGTGCATCAGAGACACAGAGGTTGCGTGATATGAACACCGTAG 1113  
Db CGCGGATTTGGTGCATCAGAGACACAGAGGTTGCGTGATATGAACACCGTAG 360  
1114 CTGCTGCACAGTCGACAGTGTGCGGCTCTCATCTGCGTGGCTGAGCTG 1173  
Db CTGCTGCACAGTCGACAGTGTGCGGCTCTCATCTGCGTGGCTGAGCTG 420  
1174 CAGACGTGTGGGGAGAGCGTCCCTTCAACCAACGCTCACAGGCTGACGCC 1233  
Db CAGACGTGTGGGGAGAGCGTCCCTTCAACCAACGCTCACAGGCTGACGCC 480  
1234 GCTTTTGTGTTTCATCACTCAACCCCTGAGGCGTGTGCGGCTGTCTCATTTGGAG 1293  
Db GCTTTTGTGTTTCATCACTCAACCCCTGAGGCGTGTGCGGCTGTCTCATTTGGAG 540  
1294 GAAAGAGTTGTCTGAGAGCGCTTTCGCGTATGCTGTGAGAGGAGTGTGAAG 1353  
Db GAAAGAGTTGTCTGAGAGCGCTTTCGCGTATGCTGTGAGAGGAGTGTGAAG 600  
1354 TCAGTGTGATGTGACGACCTCATGTTCAAGAGACTCGGGAGTCTGAGCTCGTCT 1413  
Db TCAGTGTGATGTGACGACCTCATGTTCAAGAGACTCGGGAGTCTGAGCTCGTCT 660  
1414 CTGTTTCAATGAATATCTTCTCCAGACAGCCCTGTGAGAGTACTTACCTTCTTC 1473  
Db CTGTTTCAATGAATATCTTCTCCAGACAGCCCTGTGAGAGTACTTACCTTCTTC 720  
1474 CACCTCACTCTCAGAGACTTCTGTCGCCCTTGTACTAGTGTGAGAGGCTGGAATC 1533  
Db CACCTCACTCTCAGAGACTTCTGTCGCCCTTGTACTAGTGTGAGAGGCTGGAATC 780  
1534 GAGCCAGTCTCTGCTCTGTGACGTGAAGAACAAAGAGTCCATGAGCTTAAACAG 1593  
Db GAGCCAGTCTCTGCTCTGTGACGTGAAGAACAAAGAGTCCATGAGCTTAAACAG 840  
1594 GAGCCAGTCTCTGCTCTGTGACGTGAAGAACAAAGAGTCCATGAGCTTAAACAG 1593  
Db GAGCCAGTCTCTGCTCTGTGACGTGAAGAACAAAGAGTCCATGAGCTTAAACAG 840  
841 GACGCTTCCATATCCTGCTTGTGATGAAGCGTTCCTTGTGAGCTGTGAGCGAA 900  
1654 GACGCTTCCATATCCTGCTTGTGATGAAGCGTTCCTTGTGAGCTGTGAGCGAA 1653  
Db GACGCTTCCATATCCTGCTTGTGATGAAGCGTTCCTTGTGAGCTGTGAGCGAA 900  
1654 GACGCTTCCATATCCTGCTTGTGATGAAGCGTTCCTTGTGAGCTGTGAGCGAA 1653  
Db GACGCTTCCATATCCTGCTTGTGATGAAGCGTTCCTTGTGAGCTGTGAGCGAA 900  
901 GACGCTTCCATATCCTGCTTGTGATGAAGCGTTCCTTGTGAGCTGTGAGCGAA 960  
1714 AAGCTTGTGACATGAGGCTCTCTGTGAGGTCAAGACCTTAATGCCACCCAGAGAC 1773  
Db AAGCTTGTGACATGAGGCTCTCTGTGAGGTCAAGACCTTAATGCCACCCAGAGAC 1020  
1774 AAGCTTGTGACATGAGGCTCTCTGTGAGGTCAAGACCTTAATGCCACCCAGAGAC 1773  
Db AAGCTTGTGACATGAGGCTCTCTGTGAGGTCAAGACCTTAATGCCACCCAGAGAC 1020  
961 AAGCTTGTGACATGAGGCTCTCTGTGAGGTCAAGACCTTAATGCCACCCAGAGAC 1020  
1774 AAGCTTGTGACATGAGGCTCTCTGTGAGGTCAAGACCTTAATGCCACCCAGAGAC 1773  
Db AAGCTTGTGACATGAGGCTCTCTGTGAGGTCAAGACCTTAATGCCACCCAGAGAC 1020  
1021 ACCCTGAGCGCTTCCACTGTCTTTTTCAGACTCAAGACCAAGATTTGTGCTGCA 1080  
1834 TTAACAGCTTCCAGAGAGTGTGCTTCCGATTAACAGAACTTGAATAGCACT 1893  
Db TTAACAGCTTCCAGAGAGTGTGCTTCCGATTAACAGAACTTGAATAGCACT 1140  
1894 TCTTTTGTCTTCCAGC 1909  
Db TCTTTTGTCTTCCAGC 1156

RESULT 13  
US-10-677-943-1  
; Sequence 1, Application US/10677943

Publication No. US20040072297A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as  
; APPLICANT: represented by the Secretary of the Department of Health and  
; APPLICANT: Human Services  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Tong, Zhi-Bin  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64790  
; CURRENT APPLICATION NUMBER: US/10/677, 943  
; PRIOR FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: 60/241,510  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR APPLICATION NUMBER: PCT/US02/09776  
; PRIOR FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; PRIOR FILING DATE: 2001-04-04  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 1157  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-677-943-1

Query Match 28.6%; Score 1152.8; DB 18; Length 1157;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1154; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

934 CAGGCTCCGGTACGAGATCATGTCCGACACAGAAAGCTGTTTTCATCATTTAGCGGT 813  
Db CAAGCTCCGGTACGAGATCATGTCCGACACAGAAAGCTGTTTTCATCATTTAGCGGT 60  
814 TTTCATGACCTGGGCTCTGTCTTCAACATGACCAAAAGCTTCCAAAGACTGGGCTGAG 873  
Db TTTCATGACCTGGGCTCTGTCTTCAACATGACCAAAAGCTTCCAAAGACTGGGCTGAG 120  
61 TTTCATGACCTGGGCTCTGTCTTCAACATGACCAAAAGCTTCCAAAGACTGGGCTGAG 120  
874 AAGCAGCTCCGTTCAACCTCATATGAGAGTCTGTGAGAAAGTCTGCTCCCTGAGTCC 933  
Db AAGCAGCTCCGTTCAACCTCATATGAGAGTCTGTGAGAAAGTCTGCTCCCTGAGTCC 180  
121 AAGCAGCTCCGTTCAACCTCATATGAGAGTCTGTGAGAAAGTCTGCTCCCTGAGTCC 180  
934 TTCTGATGTCACCGTCAGAGACGTGGGACAGAGAAGCTCAAGTCAGAGGTGTGCT 993  
Db TTCTGATGTCACCGTCAGAGACGTGGGACAGAGAAGCTCAAGTCAGAGGTGTGCT 240  
181 TTCTGATGTCACCGTCAGAGACGTGGGACAGAGAAGCTCAAGTCAGAGGTGTGCT 240  
994 CCCCGTACCTGTTAGTTAGAGAAATCTCCGGGAAACAAAGATCACTTGTCTTGG 1053  
Db CCCCGTACCTGTTAGTTAGAGAAATCTCCGGGAAACAAAGATCACTTGTCTTGG 300  
241 CCCCGTACCTGTTAGTTAGAGAAATCTCCGGGAAACAAAGATCACTTGTCTTGG 300  
1054 CGCGGATTTGGTGCATCAGAGACACAGAGGTTGCGTGATATGAACACCGTAG 1113  
Db CGCGGATTTGGTGCATCAGAGACACAGAGGTTGCGTGATATGAACACCGTAG 360  
301 CGCGGATTTGGTGCATCAGAGACACAGAGGTTGCGTGATATGAACACCGTAG 360  
1114 CTGCTGCACAGTCGACAGTGTGCGGCTCTCATCTGCGTGGCTGAGCTG 1173  
Db CTGCTGCACAGTCGACAGTGTGCGGCTCTCATCTGCGTGGCTGAGCTG 420  
361 CTGCTGCACAGTCGACAGTGTGCGGCTCTCATCTGCGTGGCTGAGCTG 420  
1174 CAGACGTGTGGGGAGAGCGTCCCTTCAACCAACGCTCACAGGCTGACGCC 1233  
Db CAGACGTGTGGGGAGAGCGTCCCTTCAACCAACGCTCACAGGCTGACGCC 480  
421 CAGACGTGTGGGGAGAGCGTCCCTTCAACCAACGCTCACAGGCTGACGCC 480  
1234 GCTTTTGTGTTTCATCACTCAACCCCTGAGGCGTGTGCGGCTGTCTCATTTGGAG 1293  
Db GCTTTTGTGTTTCATCACTCAACCCCTGAGGCGTGTGCGGCTGTCTCATTTGGAG 540  
481 GCTTTTGTGTTTCATCACTCAACCCCTGAGGCGTGTGCGGCTGTCTCATTTGGAG 540  
1294 GAAAGAGTTGTCTGAGAGCGCTTTCGCGTATGCTGTGAGAGGAGTGTGAATAGAAAG 1353  
Db GAAAGAGTTGTCTGAGAGCGCTTTCGCGTATGCTGTGAGAGGAGTGTGAATAGAAAG 600  
541 GAAAGAGTTGTCTGAGAGCGCTTTCGCGTATGCTGTGAGAGGAGTGTGAATAGAAAG 600  
1354 TCAGTGTGATGTGACGACCTCATGAGTTCAAGAGACTCGGGAGTCTGAGCTCCGTGCT 1413  
Db TCAGTGTGATGTGACGACCTCATGAGTTCAAGAGACTCGGGAGTCTGAGCTCCGTGCT 660

Qy	1414	CTGTTTCAACATGAACATCTCTTCTCCCAACAGCCACTGTGAGAGATACTAACCTCTTTC	1473
Db	661	CTGTTTCAACATGAACATCTCTTCTCCCAACAGCCACTGTGAGAGATACTAACCTCTTTC	720
Qy	1474	CACCTCACTTCCAGAGACTTGTGTCGCGCTTTGACTACGTGTAGAGGCGCTGGAAATC	1533
Db	721	CACCTCACTTCCAGAGACTTGTGTCGCGCGCTTTGACTACGTGTAGAGGCGCTGGAAATC	780
Qy	1534	GAGCCAGTCTCTGCCCTCTGTAGAGTTAGAGAAAGAAAGAGTCCATGAGCTTTAAACAG	1593
Db	781	GAGCCAGTCTCTGCCCTCTGTAGAGTTAGAGAAAGAAAGAGTCCATGAGCTTTAAACAG	840
Qy	1594	GCAGGCTTCCATATCCACTCGCTTTGTAGATGAAGCGCTTTCTGTTTGGCGCTGTGAGCGAA	1653
Db	841	GCAGGCTTCCATATCCACTCGCTTTGTAGATGAAGCGCTTTCTGTTTGGCGCTGTGAGCGAA	900
Qy	1654	GACGTAAAGAGAGCCACTGAGAGTCTCTGTGAGCTGTCCCGTCTCCCTGGGGGTGAAGAG	1713
Db	901	GACGTAAAGAGAGCCACTGAGAGTCTCTGTGAGCTGTCCCGTCTCCCTGGGGGTGAAGAG	960
Qy	1714	AAGCTTCTGACATGAGTCTCTCTGTTGGGTGAGCAGCCTTAATGCCACACCCACAGAGAAC	1773
Db	961	AAGCTTCTGACATGAGTCTCTCTGTTGGGTGAGCAGCCTTAATGCCACACCCACAGAGAAC	1020
Qy	1774	AACCTGGAAGGCTTCCACTGTCTTTTCAGAGACTCAAGCAAAAGATTGTTCGTTGGCA	1833
Db	1021	AACCTGGAAGGCTTCCACTGTCTTTTCAGAGACTCAAGCAAAAGATTGTTCGTTGGCA	1080
Qy	1834	TTAAACAGACTTCCAGAGAGTGTGCTTCCGATTAACAGAAACTGTGACTGTATAGCACT	1893
Db	1081	TTAAACAGACTTCCAGAGAGTGTGCTTCCGATTAACAGAAACTGTGACTGTATAGCACT	1140
Qy	1894	TCCTTCTGCTTCCAGC 1909	
Db	1141	TCCTTCTGCTTCCAGC 1156	

```

RESULT 14
US-10-399-443-3
; Sequence 3, Application US/10399443
; Publication No. US20040028669A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as Represented by the
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute
; APPLICANT: Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10/399,443
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1075
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-399-443-3

Query Match      18.8%; Score 756.6; DB 17; Length 1075;
Best Local Similarity 99.5%; Pred. No. 66-218;
Matches 759; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY      2550 TCTGGCCCTCAGCCCTCGTCGTCAGCACCAGAGCTTGACACACCTGTGCTCATTCACAACACAG 2609
db      1 TCTGGCCCTCAGCCCTCGTCGTCAGCACCAGAGCTTGACACACCTGTGCTCATTCACAACACAG 60

```

QY	2610	CTGTGGGGAACAAGGTGTAAATCTAATCTGTGTCATGATCCATGAGGCTTCCCACTGTAGCT	2665
Db	61	CTGTGGGGAACAAGGTGTAAATCTAATCTGTGTCATGATCCATGAGGCTTCCCACTGTAGCT	120
QY	2670	GCAGAGGCTGATGCTGATCAGTGCACCTTGGACAACGGCTGGCTGTGTCTTCTTGCACT	2725
Db	121	GCAGAGGCTGATGCTGATCAGTGCACCTTGGACAACGGCTGGCTGTGTCTTCTTGCACT	180
QY	2730	TGCGCTTATGGGTAACTCATGAGCTTGACGCACTTGACCTTAGCATGAACCTGTGGAAGA	2785
Db	181	TGCGCTTATGGGTAACTCATGAGCTTGACGCACTTGACCTTAGCATGAACCTGTGGAAGA	240
QY	2790	CAATGCGCTGAAGCTTCTGTGTCCGAGGTGATATGAGGAACCAATCTTGTATCTCCAGGACT	2845
Db	241	CAATGCGCTGAAGCTTCTGTGTCCGAGGTGATATGAGGAACCAATCTTGTATCTCCAGGACT	300
QY	2850	GAGTGTGGTAAAGTGCATCTCACCGCGCGCTGTGATAGAGTCTGTCTGTGATATCTC	2905
Db	301	GAGTGTGGTAAAGTGCATCTCACCGCGCGCTGTGATAGAGTCTGTCTGTGATATCTC	360
QY	2910	GAGAGGACGACACCTGAAAGACCTTGATCTCAGCGACAAATCCCTGGGTGACGGTGGGCT	2965
Db	361	GAGAGGACGACACCTGAAAGACCTTGATCTCAGCGACAAATCCCTGGGTGACGGTGGGCT	420
QY	2970	TGCTGCGCTGTGCGAGGGACTGAAGCAAAAGAACAGTGTCTTGACAGAGACTCGGGTTGAA	3022
Db	421	TGCTGCGCTGTGCGAGGGACTGAAGCAAAAGAACAGTGTCTTGACAGAGACTCGGGTTGAA	480
QY	3030	GGCAATGAGGACTGACTCTGATATGCTGTGAGGCACTCCCTTGGCCCTTCTCTGCAACCG	3085
Db	481	GGCAATGAGGACTGACTCTGATATGCTGTGAGGCACTCCCTTGGCCCTTCTCTGCAACCG	540
QY	3090	GCACTGTGACAGTCTTAAACCTGTGTGCAAGAAATTAATCTTCACTGCCCAAGAAATGATGAAGCT	3145
Db	541	GCACTGTGACAGTCTTAAACCTGTGTGCAAGAAATTAATCTTCACTGCCCAAGAAATGATGAAGCT	600
QY	3150	GTTGTTGGCCCTTGTCCCTGTCCACAGCTTAACCTTACAGATAATTTGGGCTGTGGAAATGGCA	3205
Db	601	GTTGTTGGCCCTTGTCCCTGTGTCCACAGCTTAACCTTACAGATAATTTGGGCTGTGGAAATGGCA	660
QY	3210	GTACCCCTGTGCAAAATTAAGGAAGCTGTGTGAGGAAGTGAAGTACTCAAGCCCCGAGTGT	3265
Db	661	GTACCCCTGTGCAAAATTAAGGAAGCTGTGTGAGGAAGTGAAGTACTCAAGCCCCGAGTGT	720
QY	3270	AATTGACGGTAGTGGCAATCTTTTGATGAAGATGACCGGAC	3312
Db	721	AATTGACGGTAGTGGCAATCTTTTGATGAAGATGACCGGAC	763

```

RESULT 15
US-10-677-943-3
; Sequence 3, Application US/10677943
; Publication No. US20040072297A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America as
; APPLICANT: represented by the Secretary of the Department of Health and
; APPLICANT: Human Services
; APPLICANT: Nelson, Lawrence
; APPLICANT: Tong, Zhi-Bin
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64790
; CURRENT APPLICATION NUMBER: US/10/677,943
; CURRENT FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US02/09776
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: PCT/US01/10981
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ. ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1075

```

TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-677-943-3

Query Match 18.8%; Score 756.6; DB 18; Length 1075;  
Best Local Similarity 99.5%; Pred. No. 66-218;  
Matches 759; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
OY 2550 TCTGGCCCTCAGCCCTCCGACGAAACCGAGACTTGCACACCTGTCCTATCCACAACAG 2609
    |||
Db 1 TCTGGCCCTCAGCCCTCCGACGAAACCGAGACTTGCACACCTGTCCTATCCACAACAG 60
OY 2610 CCTGGGGAACGAGGTGTAATCTACTGTGTCATCATGAGGCTTCCCACTGAGTCT 2669
    |||
Db 61 CCTGGGGAACGAGGTGTAATCTACTGTGTCATCATGAGGCTTCCCACTGAGTCT 120
OY 2670 GCAGAGCTGATGCTGAATCAGTGCACCTGACACGCTGGCTGGTCTTTCTTGCACT 2729
    |||
Db 121 GCAGAGCTGATGCTGAATCAGTGCACCTGACACGCTGGCTGGTCTTTCTTGCACT 180
OY 2730 TGGCCTTATGGGTAATCTATGCTGACGCACTGAGCCTTAGCATGAACCTGTGGAGA 2789
    |||
Db 181 TGGCCTTATGGGTAATCTATGCTGACGCACTGAGCCTTAGCATGAACCTGTGGAGA 240
OY 2790 CAATGCGCTGAAGCTTCTGTGCGAGGTGATGAGAAACCATCTTGTCACTCCAGAACCT 2849
    |||
Db 241 CAATGCGCTGAAGCTTCTGTGCGAGGTGATGAGAAACCATCTTGTCACTCCAGAACCT 300
OY 2850 GGAAGTGTAAAGTGTCACTTCACCGCCGCTGTGTGAGAGTCTGCTGTGTATCTC 2909
    |||
Db 301 GGAAGTGTAAAGTGTCACTTCACCGCCGCTGTGTGAGAGTCTGCTGTGTATCTC 360
OY 2910 GAGAGGCAACACCTGGAAGAGCTGGATCTCACGCAATGCCCTGGGTGAACGGTGGGT 2969
    |||
Db 361 GAGAGGCAACACCTGGAAGAGCTGGATCTCACGCAATGCCCTGGGTGAACGGTGGGT 420
OY 2970 TGCTGCGCTGTGCGAGGACTGAAGCAAAAGAAAGTGTCTGACGAGACTCGGGTTGAA 3029
    |||
Db 421 TGCTGCGCTGTGCGAGGACTGAAGCAAAAGAAAGTGTCTGACGAGACTCGGGTTGAA 480
OY 3030 GGCATGTGACTGACTTCTGATTTGCTGTGAGGCACTCTCTTGCCCTTTCTCGAACCG 3089
    |||
Db 481 GGCATGTGACTGACTTCTGATTTGCTGTGAGGCACTCTCTTGCCCTTTCTCGAACCG 540
OY 3090 GCATGTGACGACTTAAACCTGTGACAGATTAATTAATTTGATCCCAAGAAATGTAAGCT 3149
    |||
Db 541 GCATGTGACGACTTAAACCTGTGACAGATTAATTAATTTGATCCCAAGAAATGTAAGCT 600
OY 3150 GTGTTGCGCCTTTGCTGCTCCACGCTTAACCTTAAGATTAATTTGAGGCTGTGAAATGCA 3209
    |||
Db 601 GTGTTGCGCCTTTGCTGCTCCACGCTTAACCTTAAGATTAATTTGAGGCTGTGAAATGCA 660
OY 3210 GTACCTGTGCAATTAAGAAAGCTGTGAGGAAAGTGCAGTACTCAAGCCCGAGTCTGT 3269
    |||
Db 661 GTACCTGTGCAATTAAGAAAGCTGTGAGGAAAGTGCAGTACTCAAGCCCGAGTCTGT 720
OY 3270 AATTGACGTAAGTGGCATCTTTTGAATGAAGATGACCGACAC 3312
    |||
Db 721 AATTGACGTAAGTGGCATCTTTTGAATGAAGATGACCGGTAC 763
```

Search completed: July 19, 2005, 02:27:12  
Job time : 2360 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2005, 16:09:10 ; Search time 12236 Seconds  
(without alignments)  
12552.250 Million cell updates/sec

Title: US-10-066-521-5

Perfect score: 4035  
Sequence: 1 atggaagagagacaatcgct.....tcgaaggcgctgctgctta 4035

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hlc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1261	31.3	3475	3	AK087774 Mus muscu
2	723.8	11.9	791	7	CK000513 AGENCOURT
3	422.4	10.5	3359	3	BC021272 Homo sapi
4	422.4	10.5	3360	3	BC012789 Homo sapi
5	402.8	10.0	814	6	CB228956 AGENCOURT
6	381.4	9.5	776	7	CO814088 AGENCOURT
7	377.8	9.4	1913	3	AK016782 Mus muscu
8	374.4	9.3	808	7	CO804754 AGENCOURT
9	358.2	8.9	693	5	BU634350 UI-H-FIL-
10	357.2	8.9	684	5	BU630481 UI-H-FIL-
11	349.2	8.7	786	7	CO802335 AGENCOURT
12	336	8.3	792	7	CO828844 AGENCOURT
13	323.4	8.0	1119	4	BM454123 AGENCOURT
14	320	7.9	739	7	CO815646 AGENCOURT
15	313.6	7.8	642	5	BU618831 UI-H-FIL-
16	284.8	7.1	701	7	CO801395 AGENCOURT
17	282.4	7.0	817	7	CO797627 AGENCOURT
18	280.2	6.9	571	2	BB555431 BB555431
19	276.6	6.9	3577	3	AK087843 Mus muscu
20	267.2	6.6	3218	3	AK054378 Mus muscu
21	256.2	6.3	666	4	AV367637 AV367637
22	247.4	6.1	579	1	BM228619 K0264H01-
23	242	6.0	565	4	BM226108 K0222C06-
24	238.2	5.9	589	6	CA559886 K0264H01-

25	232.6	5.8	552	7	CN724433 E0879B11-
26	232.2	5.8	593	6	CA559979 K0266B01-
27	231.2	5.7	583	6	CA558845 K0250F09-
28	230.6	5.7	768	7	CO808001 AGENCOURT
29	229.8	5.7	643	7	CF913803 B0954D05-
30	224.8	5.6	636	7	CF914959 B0971H02-
31	221.4	5.5	515	4	BM227562 K0245E10-
32	221.4	5.5	4123	3	HSW800983
33	221	5.5	511	4	BM230035
34	221	5.5	512	4	BM226240
35	219.8	5.4	513	4	BM227869
36	218.6	5.4	519	4	BG071729
37	218.6	5.4	519	4	BG084574
38	217.8	5.4	559	6	CA561556
39	216.2	5.4	743	7	CK355847
40	213.8	5.3	3275	3	AK054426 Mus muscu
41	209.8	5.2	599	7	CF915731 B0984E03-
42	203.8	5.1	473	4	BM227899 K0251B09-
43	203.4	5.0	1434	3	AK014932 Mus muscu
44	201.2	5.0	1417	3	CF594199 full-leng
45	201.2	5.0	1586	3	CR607466 full-leng

#### ALIGNMENTS

RESULT 1	AK087774	3475 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK087774				
DEFINITION	Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:8330019P16 product:muse MATER protein (maternal-antigen-chat-embryos-require) protein, full insert sequence.				
ACCESSION	AK087774				
VERSION	AK087774.1	GI:26104500			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)				
JOURNAL	99279253				
PUBMED	10349636				
REFERENCE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)				
JOURNAL	2049374				
PUBMED	11042159				
REFERENCE	Shibata, K., Itoh, M., Aizawa, K., Nagoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M., Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Kashiwagi, K., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuzaki, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
PUBMED	20530913				
MEDLINE	11076861				
PUBMED	11076861				
REFERENCE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)				
JOURNAL	5				
REFERENCE					





Db 1578 GAATCAGCATTTTGTCTTCATTGAAAAACCAAGAGCATCATGAGATGAAAGAACTGA 1637  
 Qy 1599 CTTCCATATCCACTGGCTTTGATGAAAGGTTTCTTTGTTGGCTCTGAGGAAAGCT 1658  
 Db 1638 CGACATCGCCCTC---CTCGGATGAAAGGTTTCTTAATTTGGCTCATGAAACAGATAT 1694  
 Qy 1659 AAGAGGCCACTGAGAGTCTGTGGGCTGTCCCGTTCCCTGGGGGGTGAAGAAAGCT 1718  
 Db 1695 CTGGAAGACTGTGAGAGTCTGTGTAATATCCCGTATTCCAACTGTTGAGCAAGAGCT 1754  
 Qy 1719 TCTGCACTGGGTCTCTGTGTTGGGTGAGAGCTTAATGACACACCCAGAGAGACCT 1778  
 Db 1755 CCAACACTGGGTCTCTGATGCTCAGAGGTCAATGGCACACGCCAATGAGACCT 1814  
 Qy 1779 GAGACCTTCCACTGTCTTTTGAAGCTCAAGACAAAGATTGTTGCTGGCAATTAA 1838  
 Db 1815 GGATCCCTTCTATTTCTATTATGAGCTCAGAGATGAAGATTGTGGGGGGCTCTCA 1874  
 Qy 1839 CAGCTTCAAGAGGTGGCTTCGATTAACAGAACTGAGACTTGATGACTTTCT 1898  
 Db 1875 AGCTTCCAGAGAGTGGCTGTGATTAACAGAAAGATGAGCTTGAAGTCTTCTCA 1934  
 Qy 1899 CTGCTTCCAGACTGTCCGTATTTGCGAAATTCGGGTGATGTCAAGAGATCTTCC 1958  
 Db 1935 CTGTCTCAAGACTGTCAAGAACTTAAGGCATTCGGGTGATTAACAGACTCTCTC 1994  
 Qy 1959 AAGAGATAGTCCGTGAGGCAATGCTGTGCTCTCTAG--GATCGGGAATGAAC 2015  
 Db 1995 GGTATATATATCTCGAGCTGTGCTGTGTTATCTGTCGGGAGACAAATGTAAAGC 2054  
 Qy 2016 CCTATGAGAGAGAGTGGAGATTTTCTGCTCAAGCTTGGCACCCACCACTGCG 2075  
 Db 2055 CCTCTCATGAGTGGTGGGAACTCTGCTGTGTGGCTGGACCTTCGAACTTGA 2114  
 Qy 2076 GCAAGTGAAGCTGGGAGAGCATCTGACAGAGGCGGCATGAAGACCTGTGTCGA 2135  
 Db 2115 GGAAGTGAAGCTTGGGAGAGCATCTGATGAAGGCGCATGAAGATATCTGTCTGA 2174  
 Qy 2136 GCTGAGGCACTCCACTGCAAGATACAGACCTGATGTTTGAAGATGACAGATTAACC 2195  
 Db 2175 GCTGCGGAATAGTCTGCAAGATACAGAGCTGAGTTAAGATGACAGAGTGTGTC 2234  
 Qy 2196 TGTGTGAGAGCACTCTGAGAAATGCTCAAGCCATACCTTAAGATCCCTAACT 2255  
 Db 2235 TGGCTTGAAGCACTCTGTGAAGCTCTTTTGAAGATCAAACTTAAGTACTTAATCT 2294  
 Qy 2256 GGGAGGCAACCCACTGAAAGAAAGATGAAGTGGGTGTAAGCTTAAACACC 2315  
 Db 2295 AGGGAACACTCCCATGAAAGATGAATGAAGTTAGCTTCCGAAGCGCTGAACATTC 2354  
 Qy 2316 AAAATGTTGTTGAGTCTTTGAGGCTGATTTGCTGTGATTTGACCAATGCTTAACT 2375  
 Db 2355 AAAAGTCTCCGAGAGACTCTGAGGTGATTCGTGTGATTAACATCATTTGTTTGA 2414  
 Qy 2376 GAAATCTTCCAAATCTTACGACTCCCGCAGCTGAATCTCTGAGCTTGGCAGAA 2435  
 Db 2415 GATGATCTTCCAGCTTCTTAATTTCAACCAACGAGCTTAAGTCTCAGCTTGGCAAAA 2474  
 Qy 2436 CAAGGTGACAGACGAGGAGTGAATGCTCTGATGATGCTTGAAGTCTCCAGTGGC 2495  
 Db 2475 TAGAGTGGAGTAAAGATATATCTTGGGAATGCTTGAAGTCAATGTGCT 2534  
 Qy 2496 CTTGAGAAAGTGAATGAGAGTGTGGATCAAGCAGCGGTTCAGAGTCTGGC 2555  
 Db 2535 ACTGCAAAAGTTGATATGAGCAACTGTGGCTCAACCTGCAAGTGCACCTTCTGGT 2594  
 Qy 2556 CTGACCCCTTCGACCAACCGAGCTTGAACACTGTGCTTATCCAAACAGCTGGG 2615  
 Db 2595 CTGACCCCTTTCACCAACCAACTTGAACACTGTGCTTATCCAAACAGCTGGG 2654  
 Qy 2616 GAAACAGGTGAATCTACTGTGTGATCATGAGGCTTCCCACTGATGCTGACAG 2675  
 Db 2655 GACTGAAGAGTGAACAGCTGTGTGATCTTGAAGATCCAGATGTGCTCTCAAGC 2714

Qy 2676 GGTATGCTGATCAGTGCACCTGACAGCGGTGGTGTGTTTCTTGACTTGGCT 2735  
 Db 2715 GCTGATCTGATATCACTGCAACATTTGTGATGATGCTTATGCTTCTGCAATGACT 2774  
 Qy 2736 TATGGTAACTCATGCTGAGCAGCACTGAGCTTATGATGAACCTGTGGAAGCAATGG 2795  
 Db 2775 TGCAAAAGCAACAAAGCTGACCCACTGAGCTGACATGAACCCGTAAGGAGATGTGC 2834  
 Qy 2796 CGTGAAGCTTCTGTGCGAGTGTATGAGAAACCATTTGTATCTTCCAGACTTGGAGT 2855  
 Db 2835 AATGAAGCTACTGTGTGAAGCTTTAAAGAACTTACTTGTATCTTCAAGAACTGAACT 2894  
 Qy 2856 GGTAAAGTATCTCAACCGCGGTGTGAGAGTCTGTCTGTGATCTGAGAG 2915  
 Db 2895 AGTGAAGTCCAACTCAACAGAACTGTGTGAGAGACTGTGCTGTATGATCAACAAAC 2954  
 Qy 2916 CAGACACCTGAAGAGCTGTGATCTCAACGAAATGCTGTGGGTGAACGCTGGGGTGTGC 2975  
 Db 2955 CAGACCTTAAAGATTGATCTTGTGTAACAGCCCTGGGTGACAAAGAGTATTAAC 3014  
 Qy 2976 GCTGTGCGAGGACTGAAGCAAAAGAACAGTGTCTGAAGAGACTGGGTGAAGCAATG 3035  
 Db 3015 CTTGTGAGGAGACTGAAGCAAAAGTACAGCTCCCTGAGAGACTTGGGTGGGGCATG 3074  
 Qy 3036 TGAAGTCTTCTGATTTGCTGTGAGGCACTCTCTTGGCCCTTCTGCAACCGGATCT 3095  
 Db 3075 TAAAGTATCTTCAATGCTGTGAGGCAATGTGATGAGCAATCTTCAACCTCACCT 3134  
 Qy 3096 GACCACTTAAACCTGTGAGAAATTAATCTGATCCCAAGAAATGATGAAGCTGTTC 3155  
 Db 3135 GAAAGCTTAAACCTGTGAGAAATGATCTTCAATGATGAGGAGATTTGAAGCTGTCTC 3194  
 Qy 3156 GGCCTTTCCTGTCCAGCTTAACCTTAACATTAATGAGGCTGTGAAATGAGCACTACC 3215  
 Db 3195 TGCGTTCAATGCCCTGTCTCTTAACCTGGGATTAATGGCTGTGAGAGAGAGTACTA 3254  
 Qy 3216 TGTGCAATTAAGAAAGTGTGTGAGAGAGTGAAGTCAAGTCAAGCCCGAGTCTTAATGA 3275  
 Db 3255 TGCCGAGTGAAGAAAGCAAGCTGAGAGAGTGAAGTGTCAAGCCCGAGTGTGATGA 3314  
 Qy 3276 CGGTAGTGGCAATCTTTGATGAAGATGACCGAC 3312  
 Db 3315 TGTGATTTGTATGCTAGTGAAGATGACCGAAC 3351

RESULT 2  
 CR000513  
 LOCUS 791 bp mRNA linear EST 26-NOV-2003  
 DEFINITION AGENCOURT\_16368905 NIH\_MGC\_221 Homo sapiens cDNA clone  
 IMAGE:30708637 5', mRNA sequence.  
 ACCESSION CR000513  
 VERSION CR000513.1 GI:38526547  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 791)  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 AUTHORS Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgsaps-remail.nih.gov  
 Tissue Procurement: James Martin, University of Iowa  
 cDNA Library Preparation: M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNML)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNML at:

http://image.llnl.gov  
 Plate: NDAMI075 row: c column: 14  
 High quality sequence stop: 688.  
 Location/Qualifiers

# FEATURES

source

```
1..791
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30708637"
/lab_host="DH10B Tona"
/clone_id="NIH_MGC_221"
/note="Organ: mixed; Vector: pYX-Asc; Site_1: EcoRI;
Site_2: NotI; Library is oligo-dT primed and directionally
cloned. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Asc vector. Average insert size
4-5kb. Adaptors 5' (AATTCGACGAGG) 3' and 5' d
(CCTCGTCCG) 3'. 3' linker sequence - GCGGCGGTGAGAGCC T18.
Sequencing primers 3' end: T3 promoter primer 5'd
(AATACCTCTACTTAAGGA) 3'. 5' End: T7 promoter primer 5'd
(TAATGACTGACTATAGG) 3'. Library was constructed in the
laboratory of M. Bento Soares. Note: this is a NIH_MGC
Library"
```

## ORIGIN

```
Query Match      17.9%; Score 723.8; DB 7; Length 791;
Best Local Similarity 99.1%; Pred. No. 1.6e-188;
Matches 780; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

Oy 565 ACGGTTCTTGCACGGAAGTCAAGATTTGGGAATGGGCTCTAGCCAGAGATCGTG 624
Db 5 AGGGTTCTTGCACGGAAGTCAAGATTTGGGAATGGGCTCTAGCCAGAGATCGTG 64
Oy 625 CTGTGCTGGGCGCAAGTGTGACTTACCAAGGAATGTTCTCTTCTTCTCTCC 684
Db 65 CTGTGCTGGGCGCAAGTGTGACTTACCAAGGAATGTTCTCTTCTTCTCTCC 124
Oy 685 GTTAGAGATGACGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 744
Db 125 GTTAGAGATGACGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 184
Oy 745 CCAAGATCCCAAGGCTCCGGTACGAGATCATGTCCGACCAAGAAAGGCTGTTC 804
Db 185 CCAAGATCCCAAGGCTCCGGTACGAGATCATGTCCGACCAAGAAAGGCTGTTC 244
Oy 805 ATTGACGGTTTGATGATGATCTGTGGCTCTGTCTTCAACATGACACAAAGCTTGC 864
Db 245 ATTGACGGTTTGATGATGATCTGTGGCTCTGTCTTCAACATGACACAAAGCTTGC 304
Oy 865 TGGGCTGGAAGACAGCTTCCTGTTACCTCATAGCAATCTGTCTGAGAGAGTCTGCTC 924
Db 305 TGGGCTGGAAGACAGCTTCCTGTTACCTCATAGCAATCTGTCTGAGAGAGTCTGCTC 364
Oy 925 CCTGAGTCTTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 984
Db 365 CCTGAGTCTTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 424
Oy 985 GTGTGTCTCCCGTTACTTGTAGTGAAGAAATCTCCGGGGAACAAAGATCCACTTG 1044
Db 425 GTGTGTCTCCCGTTACTTGTAGTGAAGAAATCTCCGGGGAACAAAGATCCACTTG 484
Oy 1045 CTCTTGAAGCGGGGATTTGTGAGATCAAGAAAGACAAAGGTTGCGTCAATCAAG 1104
Db 485 CTCTTGAAGCGGGGATTTGTGAGATCAAGAAAGACAAAGGTTGCGTCAATCAAG 544
Oy 1105 AACCGTAGCTGCTGACAGGTCAGGTCGCGCGGCTCTCTATCTGGGTGGCC 1164
Db 545 AACCGTAGCTGCTGACAGGTCAGGTCGCGCGGCTCTCTATCTGGGTGGCC 604
Oy 1165 CTGACGTCGACGAGCTGTGTGGGGAGAGCGTGCCTTCAACAAAGCTCAAGGC 1224
```

```
Db 605 CTGACGTCGACGAGCTGTGTGGGAGAGAGCGTCCCTTCAACCAAGCTCAAGGC 664
Oy 1225 CTGACGTCGACGAGCTGTGTGGGATTCATCAGCTCAACCCCTGAGGCGGTGCCGCTGCTC 1284
Db 665 CTGACGTCGAGCTTTGTGTGT-ATCAGCTCAACCCCTGAGGCGGTGCCGCTGCTC 723
Oy 1285 AATTCGAGGAAAGAGTGTCTCTGAGAGCGCTTCTGCCATAGCTGTGAGGAGTGTG 1344
Db 724 AATTCGAGG-AGAGTGTCTCT-AGCGCTTGTCCGT-TGGTGTGGA-GAGTGTGG 779
Oy 1345 AATTCGA 1351
Db 780 AATTCGA 786
```

```
RESULT 3
BC021272
LOCUS
DEFINITION
Homo sapiens cDNA clone IMAGE:4811303, containing frame-shift errors.
ACCESSION
BC021272
VERSION
BC021272.2 GI:33878145
KEYWORDS
HTC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
```

```
REFERENCE
AUTHORS
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
Datchenko, L., Marsina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schneitz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McSwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Faney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalins, D.E.,
Schmetch, A., Schein, J.B., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3359)
DIRECTOR MGC Project.
REFERENCE
AUTHORS
Submitted (14-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Aug 19, 2003 this sequence version replaced gi:18204229.
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nigri.nih.gov
Antler, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
```

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, D., Pearson, R., Stantipop, S., Thomas, P.J., Touchman, J.W., Tsougen, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLM at: <http://image.llnl.gov>  
Series: IRM Plate: 39 Row: 9 Column: 3  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19745161  
This clone has the following problem: frame shifted.

## FEATURES

source

1..3359  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:481303"  
/issue\_type="Placenta, chorioarcinoma"  
/clone\_id="NH\_MGC\_21"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"

## ORIGIN

Query Match 10.5%; Score 422.4; DB 3; Length 3359;  
Best Local Similarity 50.2%; Pred. No. 9,5e-105;  
Matches 1261; Conservative 0; Mismatches 1206; Indels 45; Gaps 7;  
520 TTGGCTGGTCTTTTGGATTCAGACCGGTTGGGCTTCCGGCTTCGACGCTTTCGAC 579  
464 TTGGACCCCTTTTGGCTCCCAAGAACTGGGAAACAGCCAGTACATGATTTATCA 523  
580 GGAAGTCAAGAAATGGGAATCGCTCTAGCCAGAGAGATGCTGCTGGCGCA 639  
524 GGAACCAAGAAATGGAAACGACCTCTGATGAGCTGATGAGCTGCTGGCGAC 583  
640 GGTGACCTTCACAGGATTTCTCTACGCTTCTTCCCTCCGTTAGAGATGAC 699  
584 AACAAATCTTTCGGGATGATGCTCTGACAGCTTCTTCTGTCGACGAACTGAG 643  
700 CGAAGAAAGAGAGATGTCACAGATTCATCTCCAGGAGTGGCCAGACTCCAGGCT 759  
644 GAGTGGCCGCAACAGATTTGGCTGATCTTTCAGAGAGTGGCTGACCCCTGCT 703  
760 CCGGTCAGAGATCATGTCGACCAAGAGCTGTTGTTGATCATTTGACGTTTCAT 819  
704 CTAATACAGAGATGCTGCTCAACGAGAGACTCTTGTGTCATGACAGCTTCGA 763  
820 GACCTGGGCTGTCTCTCAACAT---GACACAAAGCTTCGAAAGATGGGCTGAG 876  
764 GAGCTGACGGGCGCTTGAAAGAACCCGATTCGATCTGTGTGACTGTATGAGAA 823  
877 CAGCTCCGTTTCAACCTCATACGAGTCTGTCGAGAAAGTCTGCTCCCTGAGCTTC 936  
824 CGGCGGTCAGAGTCTTCTGACAGTTCGTCGAGAAAGATCTCCCGAGGCTTC 883  
937 CTGATGTCACCGTCAGAGCGTGGCAAGAAAGCTCAAGTCAGAGTCTGCTCC 996  
884 CTGCTCATGTCATCAACCCGTCGTCGAGAGAGCTCCGGGATCAGGTGACATCA 943  
997 CGTACCTGTTAGTGAAGATCTCCGGGAAACAAAGATCACTTGTCTTTAGAGCC 1056  
944 GAAATCTACAGCCCGGGGATTCACAGAGTGTAGGTTAGTATTTCTGCTGTTTC 1003  
1057 GGGATTGGTGAATCAGAAAGACAGAGGTTGCGTGATCATGAACACGTTAGGCTG 1116  
1004 TTCAAGAGCCGAAAGAGCCATGAGAGCTTCAATCTTTTGAAGAAATGAACAGCTG 1063  
1117 CTGACCAAGTCAGAGTCCCGCGCTGCTCTCATCTGCGTCCCTGACGCTGAG 1176  
1064 TTTTCATATGCAATCCCGCTCTCTGCTGATCTGTGTACAGATCTGAAAGAG 1123  
1177 GACGTGTGGGGAAGGCTGCGCTTCGAACAAAGCTACAGGCTGACGCGCT 1236

1124 ATGCAAAAAGAAAAGACTGCGCTTGCACCTCCAGACATCACTCTGTGTATCTCT 1183  
1237 TTTGTGTTTCATCAGCTCAACCCCTGAGGCGTGTGCGCGCTGTCTCATCTGAGGA 1296  
1184 TTGCTTTTAACTGTTCACACCTGAGGTTGCCAGGCGCCGACCTCCGAACCCAGCA- 1242  
1297 AGAGTTGCTTGAAGCGCTTCTGCGGTATGCTGTGTGAGGAGTGTGAATGGAATCA 1356  
1243 -----CAAGCTGAGGCGCTGTGCTCCCTGCTGAGAGGATGTGACAGACATTT 1297  
1357 GTGTTGATGAGTGAAGACCTCATAGTTCAGAGACTCGGGAGTGTAGCTCGTCTG 1416  
1298 GAGTTTGTGAAGACGACCTCCGAGAAATGGGTTGTGACCTGACATCTCCGCTG 1357  
1417 TTTCAATGAACATCTCTTCCAGACAGCACTGTGAGAGTACTACATCTTCTTCAC 1476  
1358 CTGGGCAACAGATVACTTCTGAAAGTACGGGAGCGTGAAGCTCTACGTTCTCCAC 1417  
1477 CTGAGTCTCAGAGCTTCTGTGCGCTTGTATCTACGTGTAGAGGCTTGAATTCAG 1536  
1418 GTGTATTCAGAGATCTGTGCGCGCTGTGTTCTATTTGCTCAAGAGCC-----ACCTTG 1472  
1537 CAAGCTCTGCGCTCTGTAGCTTGAAGAAAGCAAGAGCTCATGAGCTTAAACAGCA 1596  
1473 ATCATCTTACCCAGCTGTGAGATGTGTACAGAAATCTGATGTTCAATTTTGA 1531  
1597 GCGCTTCATATTCATCTGCTTGTGATGAAGAGTTCCTTGTGCTGTCGAGCAAG 1656  
1532 GCAAGAGAGCAATGATTTTGTGGGGTGTTCATCGGCTTTTAAATTAAG 1591  
1657 GTAGAGAGCACTGAGAGTCTGTGGCTGTGCTGCTGCTGCGGTTGAAGCAAG 1716  
1592 GAACAGAGAAATCTGAGTGCCTTTTGTGCTTCCAACTGCTCCAGAGATTAAGCA 1651  
1717 CTTCGACCTGAGTCTCTGTGGGTGAGAGAGCTTATGACACCAACCCAGAGAC 1776  
1652 ATTCACAGCTCTGAAGAGCTTGAAGGAGGTGCAATCTCAGGAGAGGTGATTC 1711  
1777 CTGACGCGCTTCACTGCTTTCGAGACTCAAGCAAAAGATTTGCTGCTGACATTA 1836  
1712 TTGGGATATTTTACTGTCTCTTGAATGAGATCTGCTTGTGAGAGGCAAGT 1771  
1837 AACGCTTCAAGAAATGTGCTTCCATTAACAGAACTGAGCTGATAGCATCTTC 1896  
1772 AACCTCTCAAGAAAGCTTCACTTATTTTGAACAGTGGACTGTGCTTCTCC 1831  
1897 TTTGCGCTCAGACCTGCTCGATTTTGGGAAATTTGGGATGATGCAAGGATCTTC 1956  
1832 TACTGCTTAAATVACTGCTCAAGCTTGAAGAACTGTGTTTCCGTCAAAATGCTTT 1891  
1957 CCAAGAGATGAGTCCGCTGAGGACATGCTGTGCTCCTATGAGATCGGATTAAGC 2016  
1892 AAGAAAGAGAT-----GAAACACTCTACGTCGGATTAAGC 1930  
2017 CTGATTGAGAGAGTGAAGATTTCTGCTCATGCTTGGACCCACCAACCTGCGG 2076  
1931 CTGATCTGT-----TGGCATCATCTGCTGTGCTCAACACAGCGGACCTCAGA 1984  
2077 CAGCTGACCTGGGAGAGATCTGACAGAGCGGCAAGAAAGCTGTGTCAG 2136  
1985 GAGCTCCAGGTGCAAGAGACCTTCAGAGTCACTTTGTGACCTGTGTAAACAG 2044  
2137 CTGAGGCACTCCACCTGCAAGATACAGACCTGATGTTTGAAGATGCAAGATTAACCT 2196  
2045 CTGAGGATCTCCAGCTGTGCTTCAAGAGCTTGAAGATTAAGTTTCTTTCTGCG 2104  
2197 GGTGTGAGCACTCTGTGAGATCTGTAGGCAACGTTAACTTAAGTCTTCACTTG 2256  
2105 CAGAGTGTGCTCTTGAAGTGTCTTTATACAGCAGACTGAAATATCTGAGCTTC 2164  
2257 GAGGCAACCCACTGAAAGAAAGATGTAAAGATGCGTGTGAAGCTTAAACACCA 2316

```

Db 2165 ACCCTCAGCAAACTCTCTGATGATGATCAGGTCCTCTGTATGACCTTGAACCTACCA 2224
Qy 2317 AAATTTGTTGAGATCTTTGAGCGTGGATGCTGTGATTTGACCATGCTGTATACCTG 2376
Db 2225 GCAGGCAACGTCMAAGACTAGCGTGGTAAATTTGACCTCTCACCCTCATGATTTGTAA 2284
Qy 2377 AAGATCTCCAAATCTTACGACCTCCCAAGCCTGAAATCTTGAAGCTGGCAGAAAC 2436
Db 2285 GTCCCTGCTGGCCTTCTAACCAACAAGAAGTGAAGTATCTGAATGATCTCTGCAAC 2344
Qy 2437 AAGGAGACAGACAGAGAGTATGCTCTCAGTATGCTTGAAGTCTCCAGTGGCC 2496
Db 2345 CAGTT--AGACACAGGCGTGCCTTTTGTGTGAAGCCCTGTGACGCCAGACACGCTC 2401
Qy 2497 CTGCAAGCTGATGATCTGAGGATCTGTGGATCAACAGCCAGGTTGGCAGAGTGTGACC 2556
Db 2492 CTGATATCTGATGATTTGGCTTTTCTGCACTCTCAGCGAGCAGTGTGGGAATACATCTCT 2461
Qy 2557 TCAGCCTCTGACAGACCGGAGCTTGAACACCTGTGCTTATCCACAGACCTGAGG 2616
Db 2462 GAAATGCTTCTGCGTAAACAAGAGCGTGCCTATCTAGACCTCAGTGCATGCTCTGAAG 2521
Qy 2617 AACCAAGGTAAATCTAATCTGTGTGATCAATGAGGCTTCCCACTGATGTGTGAGAG 2676
Db 2522 GACCAAGGACTGAAAACCTCTCTGCGAGGCTTGAACAATCCGAGTGTGCTGATTTCA 2581
Qy 2677 CTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2736
Db 2582 CTGCTTTTGTAAATGTTTATATCATCTGCTGTGCTGTGAAGACTGCTGCTCTCTC 2641
Qy 2737 ATGGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2796
Db 2642 ATCAGCATCAAAAACCTGAATTTCTGCAAAATGGGTGCAATGAATGAGATGTGGGT 2701
Qy 2797 GTGAAGCTTCTGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 2856
Db 2702 GTGAGCTGTGTGTGCGGCTCTGTGACGATGATGATGATGATGATGATGATGATGATG 2761
Qy 2857 GTAAAGTCAATCTCAACCGCGCGCTGTGATGATGATGATGATGATGATGATGATGATG 2916
Db 2762 GAAATGTTGGTTTACGACGACCTGCTGTAAAGATCTGCGCTCTGTTCTCACTGCGACT 2821
Qy 2917 AGACACCTGAAAGCTGATCTCAACGCAATGCTGCTGCTGATGATGATGATGATGATG 2976
Db 2822 AAGACCTGCAAGAGCTCAACCTGACTTGAACACTTGGACACACAGGGGTGGTTGTA 2881
Qy 2977 CTGCGGAGGAGCTGAACCAAAAGACAGTGTCTGACAGATCTGGGTGA 3028
Db 2882 CTCTGTAGGCCCTGAGACACCCAGAGTGTGCTGACAGTGTGCTGCGGCTGA 2933

```

```

RESULT 4
LOCUS BC012789 3360 bp mRNA linear HTC 12-OCT-2004
DEFINITION Homo sapiens cDNA clone IMAGE:3957172, containing frame-shift
error.
ACCESSION BC012789
VERSION BC012789.2 GI:33872518
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3360)
REFERENCE
AUTHORS Strausberg, R.L., Feingold, F.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Dietzenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stephenson, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,

```

TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REMARK  
COMMENT  
On Aug 19, 2003 this sequence version replaced gi:15215377.

Contract: MGC help desk  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Institute for Systems Biology  
http://www.systemsbiology.org

contact: amadan@systemsbiology.org  
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov  
Series: IRAL Plate: 25 Row: K Column: 23  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19745161  
This clone has the following problem: frame shifted.

## FEATURES

location/Qualifiers  
source  
1..3360  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3957172"  
/tissue\_type="placenta, choriocarcinoma"  
/clone\_id="NIH\_MGC\_21"  
/lab\_host="DH10B-R"  
/note="Vector: pOT87"

## ORIGIN

Query Match 10.5%; Score 422.4; DB 3; Length 3360;  
Best Local Similarity 50.2%; Pred. No. 9.5e-105;  
Matches 1261; Conservative 0; Mismatches 1206; Indels 45; Gaps 7;  
Qy 520 TTGGCTGTGCTTTTGAATTCAGACCGGTGGGCTTCCGGCTTCAGAGTGTCTGCAC 579  
Db 464 TTGAGCGGCTTTTGTCTCCAGAGAACTGGAAACAGCAGCTACGATATTATTCAA 523  
Qy 580 GGAAGTCAGGAATTTGGAATCGGCTTACCCAGAAAGATCGTGTGTGGCGCA 639  
Db 524 GGAACCAAGGAATTTGGAATAACAGACCTCTGATGAAGCTGATGATGCTGTGGAC 583  
Qy 640 GTGTGACTTACAGAGGAATGTTCTCTAGTCTTCTTCCCTCCCGTTAGAGATGAG 699  
Db 584 AACCAAGATCTTTCGGATAGGTTCTGTACACGTTTATTTGTCTCAGAGACTAGG 643  
Qy 700 CGAAGAGAGAGACAGTGTACAGATTCATCTCCAGAGTGGCCAGACTCCAGGCT 759  
Db 644 GAGTTGCCCGCAAGAGATTTGCTGATTTATTTCCAGAGAGTGGCTGACCCGCTGCT 703  
Qy 760 CCGGTGACGAGATGATGCTCCGACAGAAAGCTGTGTTCATCATGACGTTTCAT 819

Db	704	CCTATTAACAGAGATCGTGTCTCAACCGGAGAGACTCTTGTTCATGACAGCTTCCAA	763
Qy	820	GACCTGGGCTCTGTCTTCAACAT---GACACAAAGCTCTGCAAAAGACTGGGCTGAGAAG	876
Db	764	GAGCTGACAGGGCGGCTGTGAACGAACCCGATTTGGGATCTGTGTGTGTGAACCTTGATGAGAAA	823
Qy	877	CAGCTTCGGTTACCTCTCATGCGAGTCTGTGTGAGAAAGTCTCTCTCTGATGCTTC	936
Db	824	CGGCGGCTGACAGTCTTCTTGAGCAGTTTGTGTGAGAAAGATCTCTCCCGAGGCTTC	883
Qy	937	CTGATCTGTCAACGGTACAGAGACGTGGGACACAGAAAGCTCAAGTACAGAGTCTGTCTCC	996
Db	884	CTGCTCAATGCGCATTAACACCGGTGTGCTCCGAGAGAGCTTCGGGATTCAGGTATGACATCTCA	943
Qy	997	CGTTACCTGTATTGATTAGAGAAATCTCCGGGAAACAAAGATCACTGTGCTCTTGAGGCG	1056
Db	944	GAAATCTACAGCCCCGGGGATTCAACGAGATGTATGGTTAGTGTATTTCTGTCTTTTC	1003
Qy	1057	GGGATTTGGTACATCAGAGACACAGAGGTTCCGTGCAATCATGACACCTGAGCTG	1111
Db	1004	TTCAAAAGACCCGAAAAGAGCCATGAGAGCTTCAATCTGTATAGAAAGATGACAGCTG	1067
Qy	1117	CTGCACAGTGTCCAGGTGCCCCGCGTGGCTCTCATCTGTGGTGGCCCTGAGCTGAC	1176
Db	1064	TTTTCATATGTCAAAATCCCGCTCTCTGTGTGATCTGTGTATCAAGTCTGAAAGAGAG	1122
Qy	1177	GACGTGTGGGGAGAGGCTCGCCCCCTTCAACAAACGCTACAGGCTGACAGCCGCT	1236
Db	1124	ATGCAGAAAGAAAGAAAGACTGGGCTGTGACCTGACAGAGACCTACCTGTGTATCTCTT	1181
Qy	1237	TTTGTGTTTCATCAGCTCACCCCTCGAGGCTGTGTCCGGGCTGTCTCAATCTGAGAGA	1296
Db	1184	TTTCGCTTTTAACCTGTTCACACTGAGGGTCCGAGGGCCGACATCCGCAACCCAGCA-	1244
Qy	1297	AGAGTTGTCTGAAACGCTTCTGCCGATGTGTGTGAGAGAGTGTGAATATGAGATCA	1356
Db	1243	----CGAGCTGAAGGCCCTGTGTCTCCCTGTGTGAGAGGGATGTGTGACAGACATTT	1297
Qy	1357	GTGTTGTGATGGAGAGACCTCAATGGTTCAAGGACTCGGGAGATCTGAGGCTCGTCTGTG	1416
Db	1298	GAGTTTGTGAAGAGACCTCCGGAGAAATGGGGTGTGTGACGCTGACATCTCTTCGCTG	1357
Qy	1417	TTTCATGAAACATCTTCTCCAGACAGCCACTGTGAGAGTACTACCTTCTTTCAC	1476
Db	1358	CTGGGCAACCAAGATACCTCTGAAAGTACGGGGAGCGGTGAGAGCTCTTAAGTGTCTCCAC	1411
Qy	1477	CTCAGTCTCCAGGACTCTGTGTCGCGCTTGTATCTAGTGTAGAGGCTCGAAATTCAG	1533
Db	1418	GTGTGTATCCAGGAGTTCTGTGTCGCGCTGTGTCTTGTGTCTMAAGCC-----ACCTVG	1472
Qy	1537	CCAGCTCTCTGGCCCTGTGATGTTAGAGAAACAAAGAGGTTCATGAGGTTTAAACAGCA	1596
Db	1473	ATCATCTTACCCAGCTGTGAGATGTACAGAAATGTCTAGTCCCAATT-TTGA AAAA	1533
Qy	1597	GGCTTCCATATCCACTGTCTTTGTGATGAAAGCGTTTCTTGTGTGCTGTGACGAGAC	1656
Db	1532	GCAAGAGAGACACTTGGATTTTTTTTGGGGTGTTTCTTAACCTGGCTTTTAAATTA AAAAG	1597
Qy	1657	GTAAGAGAGCCACTGAGAGTCTGTGCTGGCTGTCCCGTTCCCTGTGGGGTGAAGCAGAG	1718
Db	1592	GAAACAAGAAAACATGAGATGCGTTTTTTTGGCTTCCAACTGTCCCAAGATTAAGCAGCA	1651
Qy	1717	CTTCGCACTGGGCTCTCTGTGTGGGTGAGCAGGCTAATGCCACACACCCAGAGACACC	1776
Db	1652	ATTCAACAGTGTCTTAAGAGCTTAAGGGAGCGTGGCAATCTCAGAGGACAGGTGATTC	1711
Qy	1777	CTGACGCGCTTCCACTGTCTTTTTCAGACTCAAGCAAAAGATTGTTCGCTTGGCATTA	1836
Db	1712	TTGGGATATTTTACCTGTCTCTTTGAAAATGAGAGATCTGCTTTGTGAAAGCAGCTG	1777
Qy	1837	AACAGCTTCCAAAGAGTGTGCTTCGATTAACGAAACCTGTGACTTGTATGATCTTCC	1896
Db	1772	AACCTCTTCCAAAGAGCTTAACCTTATATTAATTAAGCAAGTGAACCTGTGTGTCTCC	1831

Qy	1897	TTTCGCCCTCAGCACTGTCGCGATTTGGGAAAATTGGGTGAAATGTCGAAGGATCTTC	1956
Db	1832	TATGCTTAAATATCTGCTCCAGCTTGAGGAACTCTTTTTCGTTCAAAATGTCCTT	1891
Qy	1957	CCAAAGATGATGTCGCGTGAAGCATGTCTGTGTCCTCTATGGATCGGGATTAAGAC	2018
Db	1892	AAGAAAGAGAT-----GAAACAAGCTCTACGTCGGATTACAGC	1933
Qy	2017	CTCATTTGAGGAGCAGTGGAAAGATTCTGCTCCATGCTTGGCAACCAACACTGCGG	2076
Db	1931	CTCATCTGT-----TGGCATCACTGCTCTGTGCTCAACCAACGAGGGGCACTCAGA	1984
Qy	2077	CAGCTGACCTGGGACAGACATCTTGACAAGCGGGGACATGAAGACCTCTGTGCCAAG	2133
Db	1985	GAGCTCCAGGTGCAGAGACGACCTCCGCAAGTCAGACTTTGTGACCTTGGTAAACAG	2044
Qy	2137	CTGAGGACATCCACCTGCAAGTACAGCCGTAATGTTAAGAAATGACAGATTACCCCT	2196
Db	2045	CTGAGGACATCCACGCTGTGCGCTTCAGAGCTTGGAATAAATACGTTTCTTTTCTGGC	2104
Qy	2197	GGTGTGCAGACCTCTGAGAAATGCTCAATGCGCAACCGTAACCTTAAGTCCCTCAACTTG	2255
Db	2105	CAGAGTGTCTGCTCTTTGAGGTGCTTTTATCAGCCAGACTTGAATATCCTAGCTTC	2166
Qy	2257	GGAAGCACCACTTGAAAGAAAGAGATGAAGATGCGTGTGAAGCTTAAACACCCA	2318
Db	2165	ACCCCTACGAAATCTCTCGATGATGACATCAGAGTCCCTGTGATGCTTGAATACCA	2224
Qy	2317	AAATGTTTGTGAGTCTTTGAGCTGGAATGCTGTGATGACCCATGCTCTTAACTGT	2376
Db	2225	GCAGGCACATGTCAAAGACTGCGCTGTAAATGTCACTCTACCCATTGATTGTGA	2284
Qy	2377	AAGATCTCCAAATCCTTAACAGACTCCCGCAGCCTGAATCTCTGAGCCTGGACGAAG	2436
Db	2285	GTCCTTGTGGCTCTTCAACCAACAAGAGCTGAGCTATCTGAATGTATCTGCAAC	2344
Qy	2437	AAGGTGACAGACCAAGGAGTAATGCTCTCAGTATGCTTGAAGTCTCCAGTGC	2496
Db	2345	CAGTT---AGACACAGCGGTGCCCTTTTGTGTGAAGCCGTGTGAGCCGACAGACG	2401
Qy	2497	CTGCAAGAGCTGATATCTGAGAGACTGTGGCAATCACAGCCAGGTTGCTCAAGTCTG	2556
Db	2402	CTGTGATATCTGATGTTGGCTTTCTTGCACTCAGCGAGCAGTGTGCGAATACATCT	2461
Qy	2557	TCAAGCCCTCGTACAGAACCGGAGCTTGACAACCTGTGCTTATCAACAACAGCTGG	2616
Db	2462	GAAATGCTTCTGCGTAAACAGAGCGTGTCTATCTAGACTCAAGTCCAAATGCTCTG	2521
Qy	2617	AACGAAGGTGTAATCTACTGTGTGCATCCATGAGGCTTCCCACTGATCTGCAGAG	2676
Db	2522	GACCAAGGACTGAAAATCTCTCTGCAAGGCTTGAACATCCGAGCTGTGCTGATTTCA	2581
Qy	2677	CTGATGCTGAATCAGTGGCACCTGGAACAAGCTGTGGCTGTGTTTCTTGCACTTGC	2738
Db	2582	CTGTGTTTGTAAATATGTTTATCATCTGCTGTGCTGTGAAGACCTGTGCTGTCTC	2641
Qy	2737	ATGGGTAACTATGCTGTAGCGCACTGAGCTTATGATGAACCTGTGGAAGACATATGC	2796
Db	2642	ATCAGCAATCAAAACCTGMAATTTCTGMAATTTGGGTGCAATGAATCGAGATGTGG	2701
Qy	2797	GTAAGACTTCTGTGAGAGGTGATGAGAACATCTGTGATCTTCACAGGACTTGGAATG	2856
Db	2702	GTCAGCTGTTGTGTGCGGCTCTGACGATACGGAATTTGCCCTTATGAAATCTTGTGG	2761
Qy	2857	GTAAGTGTCACTCACCGCCGCGTGTGTGAGTCTGTCTGTGTATCTGAGAGC	2916
Db	2762	GAAAGATGTGGATTAAAGAGCACTGTGTTAAGATCTCGGTCTTCTACCTGACGT	2821
Qy	2917	AGACACTGAAGAGCTGTGATCTCACGGAACAATGCCCTGGGTGACGTTGGGTTGTGCG	2976
Db	2822	AAGACCTGTGACAGCTCAACCTTACCTTGAACCTTGAACCAACAGGGGTGTGTGA	2881

OY	2977	CTGTGCGAGGGACCTGGAACCAAAAGAACGTTTCTGCAGACAATCCGGTTTA	3028
Dd	2882	CTGTGTAGGCCCTTGAGACACCACAGTGTGCTTCGACAGTGTGGCTGA	2933
RESULT 5			
LOCUS	CB228956	814 bp	mRNA linear EST 10-FEB-2003
DEFINITION	AGENCOURT_11501187 NICHD_RI_OVI Macaca mulatta cDNA clone IMAGE:6884760 5', mRNA sequence.		
ACCESSION	CB228956		
VERSION	CB228956.1	GI:28280534	
KEYWORDS	EST.		
ORGANISM	Macaca mulatta (rhesus monkey)		
	Macaca mulatta		
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;		
	Cercopithecinae; Macaca.		
REFERENCE	1 (bases 1 to 814)		
AUTHORS	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Contact: Robert Strausberg, Ph.D.		
COMMENT	Email: cgapbs-remail.nih.gov Tissue Procurement: Dr. Elliot Spindel cDNA Library Preparation: CLONTECH cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: L1CMJ315 row: b column: 23 High quality sequence stop: 235. Location/Qualifiers		
FEATURES	source		
	1..814		
	/organism="Macaca mulatta"		
	/mol_type="mRNA"		
	/db_xref="taxon:9544"		
	/clone="IMAGE:6884760"		
	/tissue.type="Ovary"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_idb="NICHD_Rh_Ovi1"		
	/note="Organ: ovary; Vector: pDNR-LIB; Site_1: Sfi I; Site_2: Sfi I; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.0-4.0 kb. Tissue pooled from pre-pubertal, post pubertal sn menopausal monkeys. Constructed by Clontech. Note: this is a NICHD library."		
ORIGIN			
Query Match	10.0%; Score 402.8; DB 6; Length 814;		
Best Local Similarity	88.6%; Pred. No. 1.5e-99;		
Matches	459; Conservative 0; Mismatches 57; Indels 2; Gaps 2;		
OY	2152	TGCAAGATAACAGACCCCTGATTGTTTGAATAATGACAGATTACCCCTGTGTGACAGACTC	2211
Dd	3	TGCAAGATAACAGACCCCTGATTGTTTGAAGAACGACAGATTGCTGTGTGACAGACTC	62
OY	2212	TGGAAATATGTCATATGCCAACCCTGAACCTTAAGATATCCCACTTGGAGGACACCAACTG	2271
Dd	63	TGGAAATATGTCATATGCCAACCCTGAACCTTAAGATATCCCACTTGGAGGACACCAACTG	122
OY	2272	AAGAAAGAGATGTATAGAGATGGCGTGTAAAGCTTAAACACCCAAAATGTTTTGTGGAG	2331
Dd	123	AAGAAAGAGATGTATATATGCGGTGTAAAGCTTAAACACCCAAAATGTTTTGTGGAG	182
OY	2332	TCTTTAGAGCTGATGCTGTGTGATTTGACCCCATGCTGTTAACCTGAAGATCTCCCAATC	2391
Dd	183	TCTTTAGAGCTGATGCTGTGTGATTTGACCCCATGCTGTTAACCTGAAGATCTCCCAATC	242
OY	2392	CTTAGAGACTCCCCCAGGCTGAATATCTCTGAGCCCTGGAGAGAAACAAGTGAACGACGAC	2451
Dd	243	CTTAGAGACTCTCCCAAGCTGAAGTCTCTGATCTCTGGAGAGAAACAAGTGAACGACAT	302

QY	2452	GGAGTATATGCTCTCAGTATATGCTTTAGAGTCTCCAGAGCGCCCGCAGGAAGCTGATA	2511
Db	303	GGAGTACGCTCTCAAGTATGCTTTAGAGGCTTCCAGTGGCCCTCGAGAACTGACA	362
QY	2512	CTGG-AGGACTGTGGATCAACAGCCAGGGGTGGCAGAGTCTGGGCTTCAAGCCCTGTAG	2570
Db	363	CTGGAAAGGACTGTGGATCAATAGCCACAGGGGTGCCAAAGTCTGGCCACAACCTGTCAG	422
QY	2571	CAACCGAGCTTGGACACACCTGTGCTTATCCAAACA-CAGCTTGGGAGAGAGTGTAA	2629
Db	423	CTACCCGGAACATGACACATCTGTGCTCAATTCACAACTATGAGGAGGAAGAGAGTGA	482
QY	2630	ATCTACTGTGTCATGCATGATGAGGCTTCCCACTGTAGT	2667
Db	483	AACCGCTGTGATCATCATTTGAGGTTCCTCCATTAAGT	520
RESULT 6			
LOCUS	CO814088	776 bp	mRNA
DEFINITION	AGENCOURT.30247088 NIH_MGC_236		linear
ACCESSION	CO814088		EST
VERSION	CO814088.1		GI:51032714
KEYWORDS	EST.		
SOURCE	Mus musculus		(house mouse)
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 776)		
TITLE	NIH-MGC http://img.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Daniela S. Gerhard, Ph.D.		
	Office of Cancer Genomics		
	National Cancer Institute / NIH		
	Bldg. 31 Rm10A07 Bethesda, MD 20892		
	Email: cgs@bbs-remail.nih.gov		
	Tissue Procurement: Dr. Kathleen Horner, Stanford University		
	cDNA Library Preparation: Express Genomics		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		
	Plate: NDAm1191 row: d column: 11		
	High quality sequence stop: 724.		
FEATURES			
source	Location/Qualifiers		
	1..776		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:30936082"		
	/lab_host="DHL0B Tona"		
	/clone_lib="NIH MGC 256"		
	/note="Organ: oocyte; Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI; cDNA was primed using oligo-dT primer: 5'-pGACTGTTCTTGAATCCGACGCGCCGCCCTT25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb resulted in an average insert size of 1.2 kb. This is a primary library (normalized primary library is NIH_MGC_257) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH_MGC library"		
ORIGIN			
Query Match	9.5%;	Score 381.4;	DB 7; Length 776;
Best Local Similarity	70.3%;	Pred. No. 1.3e-93;	
Matches 511; Conservative	0;	Mismatches 216;	Indels 0; Gaps 0;
QY	2468	GTGATGCTTTAGAGTCTCCAGAGCGCCCGCAGGAAGCTGATCTGAGGAGCTGTGGCA	2527
Db	3	GGAGTACCTTAGTACCAATGTGTCTACTGCAAAAGTGAATCTGGACAACCTGTGGCC	62

QY 2528 TCACAGCCAGGGATTGCCAGAGCTGGCTGACGCCCTGCTGACGAACCGGAGCTTGACAC 2587  
 DB 63 TCACACCTGCGCAGCTGCGACCTTCTGCTGCTTCCAGCCCTTTTCAGCAACGAAACTTGACAC 122  
 QY 2588 ACCCTGCTTATCCAAACAACAGCTGGGAGAGAGTGAATCTACTGTGTGATCCA 2647  
 DB 123 ACTCTGCTGCTCAACAACAACAGCTGGGAGAGAGTGAATCTACTGTGTGATCCA 182  
 QY 2648 TGAAGCTTCCCACTGTAATCTGTCAGAGGCTGATGCTGAATGATGCAACCTGAGACCG 2707  
 DB 183 TGAGGAATCCAGAAATGCTCTTCAGAGGCTGATGCTGAATGATGCAACCTGAGATG 242  
 QY 2708 CTGGCTGCTGCTTCTGTCAGCTTGGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2767  
 DB 243 ATGCTTATGCTTCTGTCAGAGTGAATGCTGCAACAACAAGCTGACCTGAGCC 302  
 QY 2768 TTAGCATGAACCTGTCAGAGTGAATGCTGCAACAACAAGCTGACCTGAGAGAC 2827  
 DB 303 TGACCATGAACCTGTCAGAGTGAATGCTGCAACAACAAGCTGACCTGAGAGAC 362  
 QY 2828 CATCTTGTATCTCAAGACCTGTCAGAGTGAATGCTGCAACAACAAGCTGACCTGAG 2887  
 DB 363 CTACTTGTATCTCAAGACCTGTCAGAGTGAATGCTGCAACAACAAGCTGACCTGAG 422  
 QY 2888 AGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2947  
 DB 423 AGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482  
 QY 2948 ATGCTTGTATCTCAAGACCTGTCAGAGTGAATGCTGCAACAACAAGCTGACCTGAG 3007  
 DB 483 AGCTTGTATCTCAAGACCTGTCAGAGTGAATGCTGCAACAACAAGCTGACCTGAG 542  
 QY 3008 TTCTGACAGACTCGGCTGTCAGAGTGAATGCTGCAACAACAAGCTGACCTGAG 3067  
 DB 543 CCTGAGAGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602  
 QY 3068 CCTTGGCCCTTCTGTCAGAGTGAATGCTGCAACAACAAGCTGACCTGAG 3127  
 DB 603 CATTTGCCCATCTCTTTCAGACCTGTCAGAGTGAATGCTGCAACAACAAGCTGACCTG 662  
 QY 3128 GTCCCAAGAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3187  
 DB 663 GTACATCGGGAGTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722  
 QY 3188 TAATTGG 3194  
 DB 723 TAATTGG 729  
 RESULT 7  
 AK016782 1913 bp mRNA linear HTC 03-APR-2004  
 LOCUS AK016782 Mus musculus adult male testis cDNA, RIKEN full-length enriched  
 DEFINITION library, clone:4933411P06 product:maternal effect gene, full insert  
 sequence.  
 ACCESSION AK016782  
 VERSION AK016782.1 GI:12855702  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 Carninci, P., and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
 REFERENCES Normalization and subtraction of cap-trapper-selected cDNAs to

JOURNAL Prepare full-length cDNA libraries for rapid discovery of new genes  
 MEDLINE Genome Res. 10 (10), 1617-1630 (2000)  
 PUBMED 20499374  
 11042159  
 REFERENCE 3  
 SHIBATA, K., ITOH, M., AIZAWA, K., NAGAOKA, S., SASEKI, N., CARNINCI, P.,  
 KONO, H., AKIYAMA, J., NISHI, K., KITAHARA, T., TASHIRO, H., ITOH, M.,  
 SUMI, N., ISHII, Y., NAKAMURA, S., HAZAMA, M., NISHINE, T., HARADA, A.,  
 YAMAMOTO, R., MATSUMOTO, H., SAKAGUCHI, S., IKEYAMA, T., KASHIWAGI, K.,  
 FUJIWAKE, S., INOUE, K., TOGAWA, Y., IZAWA, M., OHARA, E., WATABE, M.,  
 YONEDA, Y., ISHIKAWA, T., OZAWA, K., TANAKA, T., MATSURA, S., KAWAI, J.,  
 OKAZAKI, Y., MURAMATSU, M., INOUE, Y., KIRA, A., and HAYASHIZAKI, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861  
 REFERENCE 4  
 THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE II TEAM AND THE  
 PANTOM CONSORTIUM.  
 FUNCTIONAL ANNOTATION OF A FULL-LENGTH MOUSE cDNA COLLECTION  
 JOURNAL Nature 409, 685-690 (2001)  
 REFERENCE 5  
 THE PANTOM CONSORTIUM AND THE RIKEN GENOME EXPLORATION RESEARCH  
 GROUP PHASE I & II TEAM.  
 ANALYSIS OF THE MOUSE TRANSCRIPTOME BASED ON FUNCTIONAL ANNOTATION  
 OF 60,770 FULL-LENGTH cDNAs  
 JOURNAL Nature 420, 563-573 (2002)  
 REFERENCE 6 (bases 1 to 1913)  
 ARAKAWA, T., BONO, H., CARNINCI, P., FUKUDA, S., FUKUNISHI, Y.,  
 HIRAKAWA, T., HORI, F., IMOTANI, K., ISHII, Y., ITOH, M., IZAWA, M.,  
 KASUKAWA, T., KATO, H., KAWAI, J., KOJIMA, Y., KONO, H., KANDA, M.,  
 KOYA, S., KURIHARA, C., MATSUYAMA, T., MIYAZAKI, A., NISHI, K.,  
 NOMURA, K., NUMAZAKI, R., OHNO, M., OKAZAKI, Y., OKIDO, T., OWA, C.,  
 SAITO, H., SAITO, R., SAKAI, C., SAKAI, K., SANO, H., SASEKI, D.,  
 SHIBATA, K., SHIBATA, Y., SHINGAWA, A., SHIRAKI, T., SOGABE, Y.,  
 SUZUKI, H., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TANAKA, T.,  
 TEJIMA, Y., TOYA, T., YAMAMURA, T., YASUNISHI, A., YOSHIDA, K.,  
 YOSHINO, M., MURAMATSU, M., and HAYASHIZAKI, Y.  
 Direct Submission  
 JOURNAL Submitted (10-JUN-2000) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,  
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,  
 Fax: 81-45-503-9216)  
 COMMENT Please visit our web site (http://genome.gsc.riken.jp/) for further  
 details.  
 JOURNAL cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in Riken.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues. First strand cDNA was primed with a primer  
 [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCA 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse transcriptase  
 and subsequently enriched for full-length by cap-trapper. Second  
 strand cDNA was prepared with the primer adapter of sequence [5'  
 GAGAGAGAGATTCGAGTTAATTAAATTATATCCGCCGCCCC 3']. cDNA was cleaved  
 with BamHI and XhoI. cDNA of size comprised longer than 7 kb was  
 selected before cloning. Vector: a modified pBlueScript KS(+) after  
 bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'  
 end: BamHI. Host: DH10B.  
 FEATURES  
 source location/Qualifiers  
 1..1913  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="PANTOM DB:4933411P06"  
 /db\_xref="taxon:10090"  
 /clone="4933411P06"



```
/sex="male"
/clone_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
735..1544
/note="unannotated protein product; maternal effect gene
(MGI:1345193)"
putative"
/codon_start=1
/db_xref="GI:12855703"
/translation="MKDDMKLACEALRHPKALRDLSDCGLTPASCHLVSALES
NONTLTLCLSNNSLSTEGVQOLCOPLRPECALRDLINHCNIYDADGFLALIAN
TKLTHLSLTNNVPGVGMKRLCALKEPTCKLKEIYDCCIDMDACMTITTK
HLKSLDGLNNALGKGVLTLCRGLKSSSRIPRLGICACRSLNCCESLAIISNPH
LNSLNLKNDPSTSGKRLKCSAFQCPVSNLGIISGNLISFALFIALRRN"
polya_signal
1886..1891
/note="putative"
1913
/note="putative"
ORIGIN
Query Match          9.4%; Score 377.8; DB 3; Length 1913;
Best Local Similarity 65.7%; Pred. No. 1.8e-92;
Matches 550; Conservative 0; Mismatches 287; Indels 0; Gaps 0;
QY 2381 TCTCCCAATCTTACGACCTCCCGACCTGCAATCTCTGAGCCTGGCAGAAACAAG 2440
DB 676 TCTGGAACCTCTTTTGTAGCAATCAAACTTAACTACCTCAATTTAGGAAACACCCCA 735
QY 2441 TGACAGACGAGGAGTAAAGCTCTCAGTGAAGCTTGAAGAGTCCAGTGCAGCTGCG 2500
DB 736 TGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 795
QY 2501 AGAAGCTGATATCTGAGAGCTGTGTCATCAACGCGGTTGCCAGTCTGCTGAG 2560
DB 796 AGGCTCTGAGACTGAGACCTGTGCTGCTCAACCTGCGACCTGCGACCTTCTGCTCAG 855
QY 2561 CCTCTGTGAGAACCGGACCTTGAACACCTGTGCTTCAACAACGAGCTTGGGAAACG 2620
DB 856 CCTTTTGAAGAACGAGAACTTGAACACCTGTGCTTCAACAACGAGCTTGGGAAACG 915
QY 2621 AAGGTGAATCTTACTGTGTGTCATGATGATGATGATGATGATGATGATGATGATG 2680
DB 916 AAGAGTGAACAGCTGTGTGTCATGATGATGATGATGATGATGATGATGATGATG 975
QY 2681 TGCTGAATCACTGACCACTGACCACTGACCACTGACCACTGACCACTGACCACTG 2740
DB 976 TACTGAATCACTGACCACTGACCACTGACCACTGACCACTGACCACTGACCACTG 1035
QY 2741 GTAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2800
DB 1036 ACAACGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1095
QY 2801 AGCTTCTGTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2860
DB 1096 AGCTACTGTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1155
QY 2861 AGTGTCAATCTCAACGCGCTGTGATGATGATGATGATGATGATGATGATGATGATG 2920
DB 1156 ACTGCAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1215
QY 2921 ACTTGAAGAGCTGTGATCTCAACGATGATGATGATGATGATGATGATGATGATGATG 2980
DB 1216 ACTTGAAGAGCTGTGATCTCAACGATGATGATGATGATGATGATGATGATGATGATG 1275
QY 2981 GCGAGGAGCTGAAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCA 3040
DB 1276 GCGAGGAGCTGAAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCA 1335
QY 3041 TGACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3100
DB 1336 TGACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1395
```

```
QY 3101 GTCTAACTGTGTCAGATTAATCTTCACTGATCCCAAGAAATGATGATGATGATGATG 3160
DB 1396 GCTTAAACCTGATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1455
QY 3161 TTGCTGTCCCACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3217
DB 1456 TCCAAATGCTGTCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATG 1512
RESULT 8
CO804754 808 bp mRNA linear EST 05-AUG-2004
LOCUS CO804754
DEFINITION AGNCOURT 30257031 NIH MGC 256 Mus musculus cDNA clone
IMAGE:30938137 5', mRNA sequence.
ACCESSION CO804754
VERSION CO804754.1 GI:50992934
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 808)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cga9bbs-remail.nih.gov
Tissue Procurement: Dr. Kathleen Horner, Stanford University
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.llnl.gov
Plate: NDAM196 row: j column: 02
High quality sequence stop: 742.
Location/Qualifiers
1..808
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30938137"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 256"
/note="Organ: oocyte; Vector: pExpress-1; Site 1: EcoRV;
Site 2: NotI; cDNA was primed using oligo-dT primer;
5'-pGACTGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
resulted in an average insert size of 1.2 kb. This is a
primary library (normalized primary library is NIH MGC 257)
and was constructed by Express Genomics (Frederick, MD).
Note: this is a NIH_MGC library"
ORIGIN
Query Match          9.3%; Score 374.4; DB 7; Length 808;
Best Local Similarity 70.0%; Pred. No. 1.1e-91;
Matches 504; Conservative 0; Mismatches 216; Indels 0; Gaps 0;
QY 2462 CTCACGATGATGCTTGAAGATCTCCAGTGCAGCTGACAACTGATTAATGATGAGACT 2521
DB 10 CCTTGGGAATGCTTGAAGATCTCCAGTGCAGCTGACAACTGATTAATGATGAGACT 69
QY 2522 GTGGATCAACGAGGAGTGTGAGAGTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 2581
DB 70 GTGGATCAACGAGGAGTGTGAGAGTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 129
QY 2582 TGACACACTGTGCTATTCACAAACAGCTTGGGAAAGAGGTGAATCTACTGTGTC 2641
DB 130 TGACACACTGTGCTATTCACAAACAGCTTGGGAAAGAGGTGAATCTACTGTGTC 189
```

```

QY 2642 GATCATGAGGCTTCCCACTGATGTCAGAGGCTGATGCTGAATCAGTCCACCTCG 2701
DB 190 AGTCTTGAGGATCAGATATGCTTCACAGCGGTGATGATCACTGCAACATG 249
QY 2702 ACACGCTGCTGTGTCTTCTTGTGACCTTATGGGTAACTCATGCTGAGCACC 2761
DB 250 TAGATGATGCTTATGCTTCTTCTGCAATGAGCTTGCAACACCAAGCTGACCCAC 309
QY 2762 TGAGCTTGAAGCAACCTGTGGAAGACATAGCGGTGAAGCTTCTGTGCGAGTCA 2821
DB 310 TGAGCTTGAAGCAACCTGTGGAAGACATAGCGGTGAAGCTTCTGTGCGAGTCA 369
QY 2822 GAGAACCATCTTGTATCTCCGAGACCTGAGTGTAAAGTATCATCTGACCCGCGT 2881
DB 370 AGGAACCTTATGTTACTTCAAGAACTGGAATGATGAGCTGCAACTCACAAGAACT 429
QY 2882 GCTGTGAGAGTCTGTCTGTGTGATCTGAGAGAGACACCTGAAAGCTGATCTCA 2941
DB 430 GCTGTGAGAGACCTGCTGTATGATCACAACAACCAAGCACTTAAAGTTGATCTTG 489
QY 2942 CGGACATGCTCTGCTGTGACGCTGCGGTGCTGCGCTGTGCGAGGACTGAAGAA 3001
DB 490 GTAAACAAGCTCTGAGTGAACAAGAGTCACTAACCTGTGAGGAGCTGAAGAAAGTA 549
QY 3002 ACAAGTCTTGAACAGAGCTCGGCTTGAAGGATGAGACTGATCTTGTGCTGTGAG 3061
DB 550 GCAAGTCTCTGAGAGACTTGGGTGAGGAGTGAAGTGAATTCATTCATCTGTGAGG 609
QY 3062 CACTCTCTGAGGCTCTTCTGTCACACCGCATCTGACAGCTTAAAGCTGTGAGAA 3121
DB 610 CATTTCTATGAGGCTCTTCTGCAACCTCAGCTTAAAGCTTAAAGCTGTGAGAA 669
QY 3122 ACTTCACTCCCAAGAAATGATGAAGCTGTGCTGCGCTTGTGCTGCAAGTCTAA 3181
DB 670 ACTTCACTCCCAAGAAATGATGAAGCTGTGCTGCGCTTGTGCTGCAAGTCTTAA 729

```

```

RESULT 9
B0634350/c 693 bp mRNA linear EST 23-SEP-2002
LOCUS UI-H-FL1-bdx-m-16-0-UI-g1 NCI CGAP FL1 Homo sapiens cDNA clone
DEFINITION UI-H-FL1-bdx-m-16-0-UI 3', mRNA sequence.
ACCESSION B0634350.1 GI:23301605
VERSION B0634350
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 693)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

```

```

JOURNAL COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Martin
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@iowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

```

```

FEATURES
source
1. 693
Location/Qualifiers

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FL1-bdx-m-16-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"

```

```

/lab host="DH10B (Life Technologies)"
/clone.lib="NCI CGAP FL1"
/note="Organ: Chondrosarcoma; Vector: pT73-Pac
(pharmacia) with a modified polylinker; Site 1: EcoR I;
Site 2: Not I; NCI CGAP FL1 is a normalized cDNA library
derived from a pool of mRNA obtained from 4 cell lines
from grade III chondrosarcoma tissues. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dri)18 tail. The sequence tag for this library is
GAGGCTGCTG. The cell lines were provided by Dr. James
Martin from the University of Iowa.
TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_LIB=UI-H-FL1
TAG_SEQ=GAGGCTGCTG"

```

```

ORIGIN
Query Match 8.9%; Score 358.2; DB 5; Length 693;
Best Local Similarity 98.7%; Pred. No. 3.3e-87;
Matches 371; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```

```

QY 2937 TCTCAGGACATATGCTCTGAGTGAAGTGGGTGCTGCTGTCGAGGAGCTGAAGCA 2996
DB 693 TCTCAGGACATATGCTCTGAGTGAAGTGGGTGCTGCTGTCGAGGAGCTGAAGCA 635
QY 2997 AAAGAACAGTGTCTGAGAGACTCGGCTTGAAGGATGAGTGAATCTTGTGATGCTG 3056
DB 634 AAAGAACAGTGTCTGAGAGACTCGGCTTGAAGGATGAGTGAATCTTGTGATGCTG 575
QY 3057 TGAGGACTCTCTTGGGCTTCTGTCACACCGCATCTGACAGCTTAAAGCTGTGCA 3116
DB 574 TGAGGACTCTCTTGGGCTTCTGTCACACCGCATCTGACAGCTTAAAGCTGTGCA 515
QY 3117 GAATTAATCTTCAATCCCAAGAAATGATGAAGCTGTGCTGCTGCTGCTGCTGCTG 3176
DB 514 GAATTAATCTTCAATCCCAAGAAATGATGAAGCTGTGCTGCTGCTGCTGCTGCTGCTG 455
QY 3177 TAATTAATCAATTAATGCTGTGGAATGCAAGTACCTTGTGCAATTAAGAAAGCTGCT 3236
DB 454 TAATTAATCAATTAATGCTGTGGAATGCAAGTACCTTGTGCAATTAAGAAAGCTGCT 395
QY 3237 GAGGAGATGAGCTACTCAAGCCCGAGTGTATGAGCGGTGCTGCTGCTGCTGCTGCTG 3296
DB 394 GAGGAGATGAGCTACTCAAGCCCGAGTGTATGAGCGGTGCTGCTGCTGCTGCTGCTG 335
QY 3297 TGAAGATGACCGACAC 3312
DB 334 TGAAGATGACCGGTAAC 319

```

```

RESULT 10
B0630481/c 684 bp mRNA linear EST 23-SEP-2002
LOCUS UI-H-FL0-bdx-c-22-0-UI-g1 NCI CGAP FL0 Homo sapiens cDNA clone
DEFINITION UI-H-FL0-bdx-c-22-0-UI 3', mRNA sequence.
ACCESSION B0630481
VERSION B0630481
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 684)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

```

```

JOURNAL COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Martin
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@iowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

```

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: James Martin  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, bento-soares@iowa.edu  
Seq primer: M13 FORWARD  
POLA=Yes.

## FEATURES

Location/Qualifiers  
1..684  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UT-H-FL0-bdk-c-22-0-UT"  
/tissue\_type="Cell lines"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI CGAP FL0"  
/note="Organ: Chondrosarcoma; Vector: pT7T3-Pac  
(Pharmacia) with a modified polylinker; Site 1: EcoR I;  
Site 2: Not I; NCI\_CGAP\_FL0 is a cDNA library derived from  
a pool of mRNA obtained from 4 cell lines from grade III  
chondrosarcoma tissues. The library was constructed  
according to Bonaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoR I adaptor, digested  
with Not I, and cloned directionally into pT7T3-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tag for this library is GAGTCGCTG. The cell line  
was provided by Dr James Martin from University of Iowa.  
TAG\_TISSUE=Human Chondrosarcoma Grade 3 cell line mix  
TAG\_Lib=UI-H-FL0  
TAG\_SEQ=GAGTCGCTG"

## ORIGIN

Query Match 8.9%; Score 357.2; DB 5; Length 684;  
Best Local Similarity 98.9%; Pred. No. 6.1e-87;  
Matches 370; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 2999 TCACGGAAGAGCCCTGGGTGACGGGTGGTGGCCCTGGCGAGGACTGAAGCAAA 2998  
DB 684 TCACGGAAGAGCCCTGGGTGACGGT-GGGTGGTGGCCCTGGCGAGGACTGAAGCAAA 626  
QY 2999 AGAAGAGTGTCTGACGAGACTCGGGTTGAAGGACATGACTGACTTCTGATTGCTGTG 3058  
DB 625 AGAAGAGTGTCTGACGAGACTCGGGTTGAAGGACATGACTGACTTCTGATTGCTGTG 566  
QY 3059 AGGCACTCTCTGGCCCTTTCTTGCAACCGGCATCTGACAGCTTAAACCTGGTGCA 3118  
DB 565 AGGCACTCTCTGGCCCTTTCTTGCAACCGGCATCTGACAGCTTAAACCTGGTGCA 506  
QY 3119 ATTAATCTAGTCCCAAGGAATGATGAAGCTGTGTCGGCCCTTGGCTCCCAAGCTCA 3178  
DB 505 ATTAATCTAGTCCCAAGGAATGATGAAGCTGTGTCGGCCCTTGGCTCCCAAGCTCA 446  
QY 3179 ACTTACAGATATTGGGCTGTGGAATGGACGTACCTGTGCAATAAGAAAGCTGTG 3238  
DB 445 ACTTACAGATATTGGGCTGTGGAATGGACGTACCTGTGCAATAAGAAAGCTGTG 386  
QY 3239 AGGAAGTGAAGTACTCAAGCCCGAAGTGTGTAATGACGGTAGTGGCAATCTTTTGATG 3298  
DB 385 AGGAAGTGAAGTACTCAAGCCCGAAGTGTGTAATGACGGTAGTGGCAATCTTTTGATG 326  
QY 3299 AAGATGACGCAAC 3312  
DB 325 AAGATGACGCTAC 312

## RESULT 11

CO802335 786 bp mRNA linear EST 05-AUG-2004

## LOCUS

CO802335

## DEFINITION

AGENCOURT\_30254102 NIH\_MGC\_256 Mus musculus cDNA clone

## ACCESSION

CO802335

## VERSION

CO802335.1 GI:50990515

## KEYWORDS

EST.

## SOURCE

Mus musculus (house mouse)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 (bases 1 to 786)  
NIH-MGC <http://mgs.nci.nih.gov/>.

## AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

## TITLE

Unpublished (1999)

## JOURNAL

Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgapbs-remail.nih.gov

## COMMENT

Tissue Procurement: Dr. Kathleen Horner, Stanford University  
cDNA Library Preparation: Express Genomics  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: NDAMI201 row: k column: 11  
High quality sequence stop: 505.

## FEATURES

## source

Location/Qualifiers  
1..786  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30940090"  
/lab\_host="DH10B Tona"  
/clone\_lib="NIH MGC 256"  
/note="Organ: oocyte; Vector: pExpress-1; Site 1: EcoRV;  
Site 2: NotI; cDNA was primed using oligo-dT primer;  
5'-pGACTGTCTTCAATCGGACGGCGCCCTT25-3 and cloned into  
the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb  
resulted in an average insert size of 1.2 kb. This is a  
primary library (normalized primary library is NIH MGC 257)  
and was constructed by Express Genomics (Frederick, MD).  
Note: this is a NIH\_MGC library"

## ORIGIN

Query Match 8.7%; Score 349.2; DB 7; Length 786;  
Best Local Similarity 69.3%; Pred. No. 1.1e-84;  
Matches 474; Conservative 0; Mismatches 210; Indels 0; Gaps 0;  
QY 2462 CTCTCAGTATGCTCTTGAAGTCTCCCAAGTCCGCTCCAGACTGATATCTGAGGACT 2521  
DB 1 CCTTGGGAAGCCCTTGAAGTCTCAATGTCTTCACTGCAAAAGTTGATCTGACAACT 60  
QY 2522 GTGGCATCAGCAACGCGGTTGCGAGAGTCTGAGCCCTGCTGAGCAACCGGAGCT 2581  
DB 61 GTGGCATCAGCAACGCGGTTGCGAGAGTCTGAGCCCTGCTGAGCAACCGGAGCT 120  
QY 2582 TGACACACTGTGCTATCCAAACAAGCCCTGGGGAAGGAAGTGAATCTACTGTGTC 2641  
DB 121 TGACACACTGTGCTATCCAAACAAGCCCTGGGGAAGGAAGTGAATCTACTGTGTC 180  
QY 2642 GATCCATGAGGCTTCCCACTGTAGTCTGACAGAGCTGATGCTGAATCAAGTCCCTGG 2701  
DB 181 AGTTCTGAGGAATCCAAATGTCTCTCCAGCGGCTGATCTGAATCACTGCAACATTG 240  
QY 2702 ACAGGCTGCTGTGCTTTCTTGCACTTGGCTTAAGGGTAATCACTGAGTGAACCC 2761  
DB 241 TAGATGATGCTTATGGCTTCTGGCAATGAGACTTGGCAACCAACAAAGCTGACCCACC 300  
QY 2762 TGAGCTTAGCATGAACCTGTGGAAGCAATGGCGTGAAGCTTCTGTGCGAGGTCAATGA 2821

```

Db      301 TGAGCTGACCAATGAAACCCGATAGGGAGTGGATGATGAAAGTACTGTGAAAGCTTTAA 360
Qy      2822 GAGAACATCTTGTATCTTCACAGACCTGAGATGGTAAAGTCTATCTACCGCCGCT 2881
Db      361 AGGAACCTTACTTGTATCTTCACAGACCTGAGATGGTAAAGTCTATCTACCGCCGCT 420
Qy      2882 GCTGTAGAGTCTGTCTGTGTATCTTCAGAGAGACACACCTGAAAGCTGTGATCTCA 2941
Db      421 GCTGAGAGACCTGAGCTGTATGATCAACACACACACCTTAAAGTTGGATCTTG 480
Qy      2942 CGAACAATGCTGAGTGAACGAGTGGGTTGCTGCTGTGCGAGGAGTGAACCAAGA 3001
Db      481 GTAAACAAGCCCTGGGTGACAAAGAGTCAATACCTGTGTAGAGGACCTGAAGCAAGA 540
Qy      3002 ACAGCTTTTGAACAGACCTGCGTTGAAGACATGAGTGAATCTGTATGCTGTAGG 3061
Db      541 GCAGCTCCCTGAGAGACTTGGGTGGGCACTGATGATTTCCAAATGCTGTAGAG 600
Qy      3062 CACTGCTCTGAGCCCTTCTGCAACCGGCACTGACGATCTGAACCTGTGACGAATA 3121
Db      601 CATGTGATGAGCACTCTTGCACACCTCTCACTGAACAGCTTAACTGTGAAAGATG 660
Qy      3122 ACTCAGTCCCAAGAAATGATGA 3145
Db      661 ACTTCAGTACATCGGGAATGTTGA 684

```

```

RESULT 12      792 bp  mRNA  linear  EST 30-JUN-2004
CO382844
LOCUS          AGENCOURT 26622956 NIH_MGC_253 Rattus norvegicus cDNA clone
DEFINITION    IMAGE:7303813 5', mRNA sequence.
ACCESSION     CO382844
VERSION       CO382844.1 GI:49490946
KEYWORDS      EST.
SOURCE        Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus

```

```

REFERENCE      1. (bases 1 to 792)
AUTHORS        NIH-MGC http://mgi.mgi.nhl.gov/.
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished (1999)
COMMENT        Contact: Daniela S. Gerhard, Ph.D.
                Office of Cancer Genomics
                National Cancer Institute / NIH
                Bldg. 31 Rm10A07 Bethesda, MD 20892
                Email: egabbs-r@mail.nih.gov
                Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
                College of Wisconsin
                CDNA Library Preparation: Express Genomics
                CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Agencourt Bioscience Corporation
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LLM15339 row: 9 column: 11
                High quality sequence stop: 690.
                Location/Qualifiers

```

```

FEATURES
    source
        1..792
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /db_xref="taxon:10116"
            /clone="IMAGE:7303813"
            /tissue_type="Ovary - Brown Norway Line 3 Age 8 weeks.
            tissues were snap-frozen and transferred in -70C. RNase
            free the entire procedure."
            /lab_host="DH10B Tona"
            /lab_host="NIH MGC 253"
            /note="Organ: ovary; Vector: pExpress-1; Site_1: EcorV;
            Site_2: NotI; RNA obtained from female animals at 8 wk

```

old. Tissues were snap-frozen and kept at -80C for two days before RNA extraction and purification (Tri-reagent method). cDNA was primed using oligo-dT primer: 5'-pGACTAGTCTTAGATCGGAGCGGCC(7)25-3' and cloned into the BcoRV/NotI sites of pExpress-1. Size-selection >1.25kb resulted in an average insert size of 1.5 kb. This primary library is normalized (non-normalized primary library is NIH MGC 252) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH\_MGC library."

# ORIGIN

Query Match 8.3%; Score 336; DB 7; Length 792;  
 Best Local Similarity 66.6%; Pred. No. 4.9e-81;  
 Matches 480; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```

Qy      2102 TGACAGAGCGGCGCATGAAGACCTGTGTGCAAGCTGAGCATCCACCTGCAAGATG 2161
Db      2 TGAAATGAATGTCCTCAATGAAGATTTATGCTCAACCTGAGAAATTCATCTGACATAC 61
Qy      2162 AGACCTGATGTTTGAAGAAATGACAGATTAACCTGCTGTGTGAGACACCTGTGAGAAATG 2221
Db      62 ABAATCTTAACGTTTAAAGATTCAGAGTATGTTCTGAGCTGCAATATCTGTGATGCTCC 121
Qy      2222 TCATGAGCAACCGTAACTTAAGATCCCTCAACTTGGAGGACCCACCTGAAGAGAG 2281
Db      122 TTGTTAGCAATGAGAACTTAAAGTACTCAATCTAGGAAACAGCCCATGAAGAGATG 181
Qy      2282 ATGTAAGATGCGCTGTGAAGCTTAAACACCCAAATGTTTGTGAAGTCTTGAAGC 2341
Db      182 ACATCAAGTTAGCTGTGAAGCATGAACATCCAAAGCTGTTCCCTGAGAACCTGAGGT 241
Qy      2342 TGGATTTGCTGTGATGATGACCAATGCTGTACTGAATCTCCAAATCTTACAGACT 2401
Db      242 TGGATTTCTGTGAATGAAACCTCACTGCTGTATGAGCTATCTCCAAAGCTCTTCTTCA 301
Qy      2402 CCCCAGCTGAATATCTGTAGCTGTGAGCGTGGAGGAAACAGAGTGAACAGCCAGGAGTATG 2461
Db      302 CAGACAGCTTAAAGTGTCTCAAGCTGTGCGCAGAAATGAGTGGGTTAAAGATGACAT 361
Qy      2462 CTCTGAGTATGCTCTTGAAGATCTTCCAGTGGCCCTGACAGAGTGAATGATGAGAGCT 2521
Db      362 CCTTGGGGAAGCCTTGAAGTACTTACATGTAACCTCAAGAGCTGATGAGAGCT 421
Qy      2522 GTGGATCAACAGCCAGGCTTCCAGAGTCTGCGCCCTGACAGAGTGAATGATGAGAGCT 2581
Db      422 GTGACCTATACCTGTGAGCTGCGCAGATGTAACCTTCCGCTCTCCAGCAACCGAACT 481
Qy      2582 TGACACACCTGCTATCCAAACAGCCTGGGGAAGAGGTGAATGATGATGCTGTC 2641
Db      482 TGACTACCTGTGCTGTCAACACAGCTTGGGAGCTGAAGAAATGCAACAGCTGTGTC 541
Qy      2642 GATCCATGAGGCTTCCCACTGATGCTGACAGAGCTGATGCTGAATGATGCTGCTG 2701
Db      542 AGTCTGTAAGAAACCAAGATGTGCTCCGCGGCTGATGATGATGATGATGATGATGATG 601
Qy      2702 ACAGCGCTGCTGTGCTTCTGTGACTTGGCTTAAAGGTAATGATGATGATGATGATG 2761
Db      602 TCAAGAGATGCTTACGCGCTTCTGCGCTGATGATGCTGCAACAGAAAGCTGACGACC 661
Qy      2762 TGAGCTTGAATGATGAACCTGTGAGAAACATGAGGTAAGCTTGTGCGACATGATGA 2821
Db      662 TGAGCTTGAATGATGAACCTGTGAGAAATGATGATGATGATGATGATGATGATGATG 721
Qy      2822 G 2822
Db      722 G 722

```

```

RESULT 13      1119 bp  mRNA  linear  EST 05-FEB-2002
BM454123
LOCUS          AGENCOURT 6419582 NCBI CGAP_Ov44 Mus musculus cDNA clone
DEFINITION    IMAGE:5504420 5', mRNA sequence.

```

ACCESSION BM454123  
 VERSION BM454123.1 GI:18503163  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1119)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 TITLE Contact: Robert Strausberg, Ph.D.  
 JOURNAL Email: cgapbs-remail.nih.gov  
 COMMENT Tissue Procurement: Aaron Hsueh  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov  
 Plate: LLM12145 row: 1 column: 21  
 High quality sequence start: 3  
 High quality sequence stop: 705.  
 Location/Qualifiers  
 1..1119  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5504420"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_1ib="NCI CGAP Ov44"  
 /note="Organ: ovary; PMSG-treated; Vector: pCMV-SPORT6.1; Site\_1: EcorV; Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.2 kb. Library constructed by Life Technologies. Note: this is a NCI CGAP library."  
 ORIGIN  
 Query Match 8 0%; Score 323.4; DB 4; Length 1119;  
 Best Local Similarity 62.7%; Pred. No. 1.7e-77;  
 Matches 605; Conservative 0; Mismatches 346; Indels 14; Gaps 6;  
 QY 842 ATGACACAAAGCTCGCAAAGCTGGGCTGAGAGACACCTCCGTTCAACCTCATACCA 901  
 DB 13 ATGATATACACTATCCAGAGACTGGAAGATGAACACCCATATACATCTGATGTACA 72  
 QY 902 GTCTGCTGAGGAAGGTCTCTCTCTCTGAGTCTTCTGATGTGTCACCGTCAGAGAGCTGG 961  
 DB 73 GCTCTCTGAGGAAGGTCT 132  
 QY 962 GCACAGAGAAGCTCAAGTCAGAGGTCGTCTCCCGTTACTCTTACTTATAGAGAACTCT 1021  
 DB 133 GCTTAGAAAACCTCAAGTCATGTTGTGTCCTCTCTATATACTGTGGAAGACTGT 192  
 QY 1022 CCGGGGAACAAAGATCACTTCTCTCTGAGCGGGGATTGGTGAAGATCAGAAGAC 1081  
 DB 193 CTGATCAAGAGATCTCAGCTGCTCTGAGAAATCTCCATATAGATCTGATAGAAATAC 252  
 QY 1082 AAGGTGCTGTCATCAATGAACAAACCTGAGCTCTGACCAAGTCCAGGTCGCCG 1141  
 DB 253 AAGTCTTCATCTCTGATAGAAATCAACAGCTCTTGAACCAATGCCAGGCCCTCTG 312  
 QY 1142 TGGGCTCTCATCTGCGGCTGCGCTGAGCTGCAAGAGCTGTGGGGAGAGCGTCCG 1201  
 DB 313 TGTGCTCCCTGCTGTAGAGCTTACAGCTACAAAGAACTGGGAAAAGATGCACCC 372  
 QY 1202 CTTTCAACCAACGCTCAGAGCTGACGCGCTTTTGTGTTTATGATGATCAAGCTC 1261  
 DB 373 TACCTGCAAGACTCTACCGGTTTGTATGACACCTTGTGTTTACACGCTCAGCTTGA 432  
 QY 1262 GAGGCTGATCGCGGCTGTCTCAATCTGAGAGAAAGTTGTCTTGAAGCGCTTCTGCC 1321  
 DB 433 AAAAGCTTCCAGAGCGCTCTCACTCAGGAAGAACAGATTACTTAAAGGGGTTGTGCA 492

QY 1322 GTATGCTGTGAGAGAGTGTGAATAGAAAGTCACTGTTGATGTGAGACACTCATTAG 1381  
 DB 493 TGATGGACGCTGAAGAGAGTGTGACATGAGTGTGCTTATGATGATGACCTGAAGA 552  
 QY 1382 TTCAAGACTCGGGAGCTGAGCTCCGTGCTCTGTTTCAATGAAACATCTTCTCCAG 1441  
 DB 553 ACTATAGCTTAAAGAGCTGAGATCTTGGCCCTCTTTCATGATCAATCTTCTCCAGG 612  
 QY 1442 ACAGCC--ACTGTAGAGAGTACTACACCTTCTTCCACTTCAAGTCTCAGACTTCTGTG 1498  
 DB 613 TTGGCCCAACAGGACAGATGTTATGTTTCTCCACTCAGCTCAGAGATTCTTTG 672  
 QY 1499 CCGCTTGTACTAGCTGTGTGAGGGCCCTGGAATATGACGCAAGCTCTGACCTCTGTAAG 1558  
 DB 673 CTGCTTATATATTATGTTTGAAGAGGCTGA--GGATGATATCAGATTTTGTCTCA 729  
 QY 1559 TTGAGAAACAAAGAGGTTCATGAGCTTAAACAGGAGGCTTGCATATCCATCGCTTT 1618  
 DB 730 TTGAAAACAAAGAGGATCATGAGGTGA--AGAGAACTGACGACCTGCTCTCG 786  
 QY 1619 GGATGAAGCTTTCTTGTGTTGGCTCTGAGCGAAAGCTTAAGAGGCACT--GAGGTC 1677  
 DB 787 GGATGAAGCGTTTCTTATTTGGCTCATGAAACAAGATATCTTGAAGACTCTGGAGGTT 846  
 QY 1678 CTGCTGGGCTGTCCCGTTCCCTGCGGGGTGAAGAGCTTCTGCACTGGGCTCTCTG 1737  
 DB 847 CTGTTTGAATATCCCGTATCTTCACTGT--TGACAAAGCTCCACACTGGTCTCTCTG 904  
 QY 1738 TTGGGTGACGACCTTAATGACCAACCCACAGAGACACCTGAGCGCTTCACTGTCTT 1797  
 DB 905 ATAGCTGACGAGGCGATGAGCACCGCCCAA--TGACACCTGATGATGCTTCAATTGGCTA 962  
 QY 1798 TTGCA 1802  
 DB 963 TTGCA 967  
 RESULT 14  
 LOCUS COB15646 739 bp mRNA linear EST 06-AUG-2004  
 DEFINITION AGENCOURT\_30254302 NIH\_MGC\_256 Mus musculus cDNA clone  
 ACCESSION COB15646  
 VERSION COB15646.1 GI:51035046  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 739)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 TITLE Contact: Daniela S. Gerhard, Ph.D.  
 JOURNAL Office of Cancer Genomics / NIH  
 COMMENT National Cancer Institute / NIH  
 Blog: 31 Rm10A07 Bethesda, MD 20892  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Dr. Kathleen Horner, Stanford University  
 cDNA Library Preparation: ExpreS Genomics  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov  
 Plate: NDAM1201 row: c column: 24  
 High quality sequence stop: 650.  
 Location/Qualifiers  
 1..739  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"

```

/clone="IMAGE:30939911"
/lab host="DH10B Tona"
/clone_lib="NIH_MGC_256"
/notes="Organ: oocyte; Vector: pExpress-1; Site_1: EcorV;
Site_2: NotI; cDNA was primed using oligo-dt primer:
5'-pGACTGATCTAGATCGAGCGGCCGCTT(25-3' and cloned into
the EcorV/NotI sites of pExpress-1. Site-selection >0.5 kb
resulted in an average insert size of 1.2 kb. This is a
primary/library (normalized primary library is NIH_MGC_257)
and was constructed by Express Genomics (Frederick, MD).
Note: this is a NIH_MGC library"

```

## ORIGIN

```

Query Match 7.9%; Score 320; DB 7; Length 739;
Best Local Similarity 69.5%; Pred. No. 1.3e-76;
Matches 448; Conservative 0; Mismatches 196; Indels 1; Gaps 1;

QY 2464 CTGAGTGAATGCTTGAAGATCTCCAGTGCCTGCGAGAACTGATCTGAGAGACTGT 2523
DB 1 CTTGGGAATGCTTGAAGATCTCCAGTGCCTGCGAGAACTGATCTGAGAGACTGT 60
QY 2524 GGCATCAACGACGCGGTTGCCAGAGTCTGCGCTCAGCCCTGCTGACGACCGAGCTTG 2583
DB 61 GGCCTCAACGACGCGGTTGCCAGAGTCTGCGCTCAGCCCTGCTGACGACCGAGCTTG 120
QY 2584 ACACACCTGTGCTTCCAAACACGCTGGGGAAGAGTGAATCTACTGTGTGCA 2643
DB 121 ACACACCTGTGCTTCCAAACACGCTGGGGAAGAGTGAATCTACTGTGTGCA 180
QY 2644 TCCATGAGGCTTCCCACTGATGTGAGAGGCTGATCTGATCAATGAGCCACTGAC 2703
DB 181 TTCCTGAGGAATTCAGAAATGTCTCTCCAGCGCTGATCTGATCAATGAGCCACTGAC 240
QY 2704 ACGGCTGCTGTGTTTCTTTCACCTTGGCTTATGAGGTAATGCTGAGCGACTG 2763
DB 241 GATGATGCTTATGCTTCTGCGCAATGAGACTTGCACCAAGCTGACCACTG 300
QY 2764 AGCCTTAGCATTAACCTCTGTGAGAGCAATGCGGTGAGAGCTTGTGCGAGCTCATGAGA 2823
DB 301 AGCCTTAGCATTAACCTCTGTGAGAGCAATGCGGTGAGAGCTTGTGCGAGCTCATGAGA 360
QY 2824 GAACCATCTTGTGATCTCCAGAGCTGAGAGTGTAAAGTGCATCTCACTCCGCGGTCG 2883
DB 321 GAACCATCTTGTGATCTCCAGAGCTGAGAGTGTAAAGTGCATCTCACTCCGCGGTCG 420
QY 2884 TGTGAGAGTGTCTGCTGTGATCTGAGAGCAACACTGAGAGCTTGTGATCTGAG 2943
DB 421 TGTGAGAGTGTCTGCTGTGATCTGAGAGCAACACTGAGAGCTTGTGATCTGAG 480
QY 2944 GACATGCTGCTGTGAGAGCTGAGAGTGTGCTGCTGTGCGAGAGCTGAGAGCAAAAGAC 3003
DB 481 AACAACGCTGCTGTGAGAGCTGAGAGTGTGCTGCTGTGCGAGAGCTGAGAGCAAAAGAC 540
QY 3004 AGTGTCTGAGAGAGCTGGGTTGAAGAGTGTGAGAGCTTGTGATCTGAGAGCA 3063
DB 541 AGTGTCTGAGAGAGCTGGGTTGAAGAGTGTGAGAGCTTGTGATCTGAGAGCA 600
QY 3064 CTCTCTGCTGCTTCTTCTGCAACCGGATCTGACAGCTTAAC 3108
DB 601 TGTGATTTGGGCTCTTTG-AACTCTCACTGAAAGAGCTTAAC 644

```

## RESULT 15

```

BUE18831/c 642 bp mRNA linear EST 23-SEP-2002
LOCUS BUE18831 UI-H-FH1-BFL-1-19-0-UI-81 NCI CGAP FH1 Homo sapiens cDNA clone
DEFINITION UI-H-FH1-BFL-1-19-0-UI-81 mRNA sequence.
ACCESSION BUE18831.1 GI:23285046
VERSION BUE18831.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

## REFERENCE

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 642)  
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CCAP),  
Tumor Gene Index  
Unpublished (1997)

## JOURNAL

Contact: Robert Strausberg, Ph.D.

## COMMENT

Email: ccgaps-remail.nih.gov  
Tissue Procurement: James Martin  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, bento-soares@uiowa.edu  
Seq primer: M13 FORWARD  
POLYA=Yes.

## FEATURES

Location/Qualifiers  
1..642  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-FH1-BFL-1-19-0-UI"  
/cissue\_type="Cell Line"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"

/clone\_lib="NCI CGAP FH1"  
/note="Organ: Chondrosarcoma; Vector: pRTT3-Pac  
(Pharmacia) with a modified polylinker; Site\_1: Ecor I;  
Site\_2: Not I; NCI CGAP FH1 is a normalized cDNA library  
obtained from a cell line derived from grade I  
chondrosarcoma tissue. The library was constructed and  
normalized according to Bonaldi, Lennan and Soares, Genome  
Research, 6:791-806, 1996. First strand cDNA synthesis was  
primed with an oligo-dT primer containing a Not I site.  
Double stranded cDNA was ligated to an Ecor I adaptor,  
digested with Not I, and cloned directionally into  
pRTT3-Pac vector. The oligonucleotide used to prime the  
synthesis of first-strand cDNA contains a library tag  
sequence that is located between the Not I site and the  
(dT)18 tail. The sequence tag for this library is  
AGAAATCCGGC. The cell line was provided by Dr. James Martin  
from the University of Iowa.  
TAG\_TISSUE=Human Chondrosarcoma Cell Line C88 - Grade 1  
Chondrosarcoma  
TAG\_LIB=UI-H-FH1  
TAG\_SEQ=AGAAATCCGGC"

## ORIGIN

```

Query Match 7.8%; Score 313.6; DB 5; Length 642;
Best Local Similarity 98.5%; Pred. No. 7.3e-75;
Matches 327; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 2981 GCGAGGAGCTGAAGAAAGAAAGAGTGTGAGAGAGCTGGGTTGAAGGATGTGAC 3040
DB 642 GCGAGGAGCTGAAGAAAGAAAGAGTGTGAGAGAGCTGGGTTGAAGGATGTGAC 584
QY 3041 TGACTTCTGATTTGCTGTGAGGACCTCTCTGCGCTTCTTCTGCAACCGGATCTGACCA 3100
DB 583 TGACTTCTGATTTGCTGTGAGGACCTCTCTGCGCTTCTTCTGCAACCGGATCTGACCA 524
QY 3101 GTCTTAACCTGTGTGAGAGTAATCTTCACTGCTCCCAAGAGATGATGAAGCTGTGTGGCT 3160
DB 523 GTCTTAACCTGTGTGAGAGTAATCTTCACTGCTCCCAAGAGATGATGAAGCTGTGTGGCT 464
QY 3161 TTGCTGTGCGACGCTTAACCTTAAGATTAATGAGGAGTGAAGGAGTGAAGGAGTGAAG 3220
DB 463 TTGCTGTGCGACGCTTAACCTTAAGATTAATGAGGAGTGAAGGAGTGAAGGAGTGAAG 404
QY 3221 AATTAAGAGAGCTGTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3280
DB 403 AATTAAGAGAGCTGTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 344
QY 3281 GTTGGCATTTCTTTGATGAAGATGACCGACAC 3312

```

Db 343 TTTGGCATTCTTTGATGAGATGACCGGTAC 312

Search completed: July 19, 2005, 01:37:18  
Job time : 12247 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2005, 12:03:24 ; Search time 32 Seconds  
(without alignments)  
3135.261 Million cell updates/sec

Title: US-10-066-521-6  
Perfect score: 1344  
Sequence: 1 MEGDKSLTFPSYGLQWCLYE.....DDHSGVMSLGAAGLEGIVS 1344

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size: 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*\n2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*\n3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*\n4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*\n5: /cgn2\_6/ptodata/1/1aa/PCUTS.COMB.pep:\*\n6: /cgn2\_6/ptodata/1/1aa/backfilest.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	0.7	209	US-09-388-221B-19	Sequence 19, Appl
2	9	0.7	1399	US-09-388-221B-4	Sequence 4, Appl
3	9	0.7	1424	US-09-388-221B-12	Sequence 12, Appl
4	9	0.7	1443	US-09-388-221B-6	Sequence 6, Appl
5	9	0.7	1454	US-09-388-221B-10	Sequence 10, Appl
6	9	0.7	1473	US-09-388-221B-2	Sequence 2, Appl
7	8	0.6	150	US-09-270-767-32603	Sequence 32603, A
8	8	0.6	150	US-09-270-767-47820	Sequence 47820, A
9	8	0.6	296	US-09-252-991A-22223	Sequence 22223, A
10	8	0.6	321	US-09-902-540-10274	Sequence 10274, A
11	8	0.6	376	US-09-902-540-9957	Sequence 9957, Ap
12	8	0.6	388	US-09-248-796A-17824	Sequence 17824, A
13	8	0.6	661	US-09-371-338-7	Sequence 7, Appl
14	7	0.5	59	US-09-270-767-57972	Sequence 57972, A
15	7	0.5	60	US-09-543-681A-4574	Sequence 4574, Ap
16	7	0.5	60	US-09-513-999C-7527	Sequence 7527, Ap
17	7	0.5	63	US-09-513-999C-6481	Sequence 6481, Ap
18	7	0.5	94	US-09-270-767-59952	Sequence 59952, A
19	7	0.5	108	US-09-216-393B-4	Sequence 4, Appl
20	7	0.5	110	US-09-543-681A-6458	Sequence 6458, Ap
21	7	0.5	111	US-09-621-976-7083	Sequence 7083, Ap
22	7	0.5	117	US-10-101-464A-782	Sequence 782, App
23	7	0.5	118	US-09-270-767-41140	Sequence 41140, A
24	7	0.5	118	US-09-270-767-56356	Sequence 56356, A
25	7	0.5	126	US-09-640-211A-2108	Sequence 2108, App
26	7	0.5	128	US-09-732-210-647	Sequence 647, App
27	7	0.5	128	US-09-513-999C-7418	Sequence 7418, App

28	7	0.5	130	US-10-101-464A-780	Sequence 780, App
29	7	0.5	151	US-09-270-767-33903	Sequence 33903, A
30	7	0.5	174	US-09-270-767-44505	Sequence 44505, A
31	7	0.5	176	US-09-270-767-32581	Sequence 32581, A
32	7	0.5	176	US-09-270-767-47798	Sequence 47798, A
33	7	0.5	180	US-09-489-039A-9331	Sequence 9331, Ap
34	7	0.5	183	US-09-489-039A-7626	Sequence 7626, Ap
35	7	0.5	187	US-09-716-129-58	Sequence 58, Appl
36	7	0.5	189	US-08-779-764A-28	Sequence 28, Appl
37	7	0.5	189	US-08-779-764A-29	Sequence 29, Appl
38	7	0.5	189	US-08-779-764A-30	Sequence 30, Appl
39	7	0.5	189	US-09-563-456-28	Sequence 28, Appl
40	7	0.5	189	US-09-563-456-29	Sequence 29, Appl
41	7	0.5	189	US-09-563-456-30	Sequence 30, Appl
42	7	0.5	202	US-09-270-767-42656	Sequence 42656, A
43	7	0.5	220	US-09-902-540-12909	Sequence 12909, A
44	7	0.5	221	US-09-489-039A-12495	Sequence 12495, A
45	7	0.5	231	US-09-489-039A-8999	Sequence 8999, Ap

## ALIGNMENTS

RESULT 1  
US-09-388-221B-19  
Sequence 19, Application US/09388221B  
Patent No. 6818750  
GENERAL INFORMATION:  
APPLICANT: Reed, John C.  
TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation  
FILE REFERENCE: P-LJ 3650  
CURRENT APPLICATION NUMBER: US/09/388,221B  
CURRENT FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 19  
LENGTH: 209  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-388-221B-19

Query Match 0.7%; Score 9; DB 4; Length 209;  
Best Local Similarity 100.0%; Pred. No. 4.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 260 SRPERLLFI 268  
Db 72 SRPERLLFI 80

RESULT 2  
US-09-388-221B-4  
Sequence 4, Application US/09388221B  
Patent No. 6818750  
GENERAL INFORMATION:  
APPLICANT: Reed, John C.  
TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation  
FILE REFERENCE: P-LJ 3650  
CURRENT APPLICATION NUMBER: US/09/388,221B  
CURRENT FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 1399  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-388-221B-4

Query Match 0.7%; Score 9; DB 4; Length 1399;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 260 SRPERLLFI 268

Db 400 SRPERLLFI 408

RESULT 3  
US-09-388-221B-12  
Sequence 12, Application US/09388221B

Patent No. 6818750  
GENERAL INFORMATION:  
APPLICANT: Reed, John C.  
TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation  
FILE REFERENCE: P-LJ 3650  
CURRENT APPLICATION NUMBER: US/09/388,221B  
CURRENT FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 12  
LENGTH: 1424  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic Construct  
US-09-388-221B-12

Query Match 0.7%; Score 9; DB 4; Length 1424;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 SRPERLLFI 268  
Db 400 SRPERLLFI 408

RESULT 4  
US-09-388-221B-6  
Sequence 6, Application US/09388221B

Patent No. 6818750  
GENERAL INFORMATION:  
APPLICANT: Reed, John C.  
TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation  
FILE REFERENCE: P-LJ 3650  
CURRENT APPLICATION NUMBER: US/09/388,221B  
CURRENT FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 1443  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-388-221B-6

Query Match 0.7%; Score 9; DB 4; Length 1443;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 SRPERLLFI 268  
Db 400 SRPERLLFI 408

RESULT 5  
US-09-388-221B-10  
Sequence 10, Application US/09388221B

Patent No. 6818750  
GENERAL INFORMATION:  
APPLICANT: Reed, John C.  
TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation  
FILE REFERENCE: P-LJ 3650  
CURRENT APPLICATION NUMBER: US/09/388,221B  
CURRENT FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 10

LENGTH: 1454.  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic Construct  
US-09-388-221B-10

Query Match 0.7%; Score 9; DB 4; Length 1454;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 SRPERLLFI 268  
Db 400 SRPERLLFI 408

RESULT 6  
US-09-388-221B-2  
Sequence 2, Application US/09388221B

Patent No. 6818750  
GENERAL INFORMATION:  
APPLICANT: Reed, John C.  
TITLE OF INVENTION: No. 6818750e1 Card Proteins Involved in Cell Death Regulation  
FILE REFERENCE: P-LJ 3650  
CURRENT APPLICATION NUMBER: US/09/388,221B  
CURRENT FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1473  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-388-221B-2

Query Match 0.7%; Score 9; DB 4; Length 1473;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 SRPERLLFI 268  
Db 400 SRPERLLFI 408

RESULT 7  
US-09-270-767-32603  
Sequence 32603, Application US/09270767

Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 32603  
LENGTH: 150  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-270-767-32603

Query Match 0.6%; Score 8; DB 4; Length 150;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1197 AELEGLGS 1204  
Db 100 AELEGLGS 107

RESULT 8  
US-09-270-767-47820  
Sequence 47820, Application US/09270767

```
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47820
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47820

Query Match          0.6%; Score 8; DB 4; Length 150;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1197 ALEGLGS 1204
      |||||
Db      100 ALEGLGS 107

RESULT 9
US-09-252-991A-22293
; Sequence 22293, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22293
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22293

Query Match          0.6%; Score 8; DB 4; Length 296;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      849 QSLASALV 856
      |||||
Db      106 QSLASALV 113

RESULT 10
US-09-902-540-10274
; Sequence 10274, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10274
; LENGTH: 321
```

```
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10274

Query Match          0.6%; Score 8; DB 4; Length 321;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      194 GKSGIGKS 201
      |||||
Db      158 GKSGIGKS 165

RESULT 11
US-09-902-540-9957
; Sequence 9957, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9957
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-9957

Query Match          0.6%; Score 8; DB 4; Length 376;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      554 RRPLEVLL 561
      |||||
Db      120 RRPLEVLL 127

RESULT 12
US-09-248-796A-17824
; Sequence 17824, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANT
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17824
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17824

Query Match          0.6%; Score 8; DB 4; Length 388;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      302 LIRKVLIP 309
      |||||
Db      47 LIRKVLIP 54
```

RESULT 13  
US-09-371-338-7  
; Sequence 7, Application US/09371338  
; Patent No. 6613959  
; GENERAL INFORMATION:  
; APPLICANT: Sheen, Jen  
; APPLICANT: Koyun, Yelena V.  
; APPLICANT: Chiu, Wan-ling  
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING A MAPKKK  
; FILE REFERENCE: 00786/366002  
; CURRENT APPLICATION NUMBER: US/09/371,338  
; CURRENT FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/095,938  
; PRIOR FILING DATE: 1998-08-10  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO: 7  
; LENGTH: 661  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-371-338-7

Query Match 0.5%; Score 7; DB 4; Length 661;  
Best Local Similarity 100.0%; Pred. No. 99;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 274 DLGSVLNN 281  
|||||  
Db 337 DLGSVLNN 344

RESULT 14  
US-09-270-767-57972  
; Sequence 57972, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 57972  
; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-57972

Query Match 0.5%; Score 7; DB 4; Length 59;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 NNILLPD 481  
|||||  
Db 53 NNILLPD 59

RESULT 15  
US-09-543-681A-4574  
; Sequence 4574, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO: 4574  
; LENGTH: 60  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-4574

Query Match 0.5%; Score 7; DB 4; Length 60;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 804 LKSLSLA 810  
|||||  
Db 44 LKSLSLA 50

Search completed: July 19, 2005, 12:19:48  
Job time : 35 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2005, 12:11:18 ; Search time 124 Seconds

(without alignments)  
4207.459 Million cell updates/sec

Title: US-10-066-521-6

Perfect score: 1344

Sequence: 1 MEGDKSLTFSSYGLQWCLYE.....DDHSGVMSLGNAGLEGLVS 1344

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1736639 seqs, 388188149 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1736639

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*\n2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*\n3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*\n4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*\n5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*\n6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*\n7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*\n8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*\n9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*\n10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*\n11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*\n12: /cgn2\_6/ptodata/1/pubpaa/US09C\_NEW\_PUB.pep:\*\n13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*\n14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*\n15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*\n16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*\n17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep:\*\n18: /cgn2\_6/ptodata/1/pubpaa/US10F\_NEW\_PUB.pep:\*\n19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep:\*\n20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*\n21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*\n22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1344	100.0	1344	US-10-124-498-6	Sequence 6, Appli
2	1344	100.0	1344	US-10-066-521-6	Sequence 6, Appli
3	1006	74.9	1162	US-10-216-642-1	Sequence 2, Appli
4	1006	74.9	1162	US-10-416-642-1	Sequence 1, Appli
5	988	73.5	1143	US-10-216-645-4	Sequence 4, Appli
6	968	72.0	1033	US-10-092-900A-348	Sequence 348, App
7	602	44.8	1200	US-10-399-443-24	Sequence 24, Appli
8	602	44.8	1200	US-10-677-943-24	Sequence 24, Appli
9	602	44.8	1200	US-10-860-761-4	Sequence 96, Appli
10	544	40.5	682	US-10-407-866-96	Sequence 96, Appli
11	158	11.8	385	US-10-399-443-2	Sequence 2, Appli

12	158	11.8	385	15	US-10-677-943-2	Sequence 2, Appli
13	149	11.1	258	15	US-10-399-443-4	Sequence 4, Appli
14	149	11.1	258	15	US-10-677-943-4	Sequence 4, Appli
15	12	0.9	190	10	US-09-965-621-63	Sequence 63, Appli
16	12	0.9	190	15	US-10-407-866-63	Sequence 63, Appli
17	12	0.9	190	16	US-10-781-294-63	Sequence 6, Appli
18	12	0.9	321	15	US-10-407-866-116	Sequence 116, App
19	12	0.9	449	14	US-10-028-374-2	Sequence 2, Appli
20	12	0.9	449	14	US-10-183-770-2	Sequence 2, Appli
21	12	0.9	635	15	US-10-407-866-90	Sequence 90, Appli
22	12	0.9	1004	15	US-10-108-260A-3161	Sequence 3161, Ap
23	12	0.9	1027	15	US-10-407-866-68	Sequence 68, Appli
24	12	0.9	1035	10	US-09-965-621-24	Sequence 24, Appli
25	12	0.9	1035	15	US-10-407-866-24	Sequence 24, Appli
26	12	0.9	1035	16	US-10-781-294-24	Sequence 24, Appli
27	12	0.9	1061	14	US-10-124-498-18	Sequence 18, Appli
28	12	0.9	1061	14	US-10-066-521-18	Sequence 5, Appli
29	12	0.9	1111	14	US-10-216-645-5	Sequence 6, Appli
30	12	0.9	1111	15	US-10-399-443-6	Sequence 6, Appli
31	12	0.9	1111	15	US-10-677-943-6	Sequence 2387, A
32	9	0.7	80	14	US-10-029-386-29387	Sequence 257211,
33	9	0.7	87	16	US-10-425-115-257211	Sequence 34582, A
34	9	0.7	162	9	US-09-864-761-34582	Sequence 4432, Ap
35	9	0.7	192	10	US-09-764-891-4432	Sequence 13890, A
36	9	0.7	224	17	US-10-732-923-13890	Sequence 223227,
37	9	0.7	284	15	US-10-424-599-223227	Sequence 114, App
38	9	0.7	303	14	US-10-306-762-114	Sequence 92, Appli
39	9	0.7	764	15	US-10-407-866-92	Sequence 4, Appli
40	9	0.7	1399	9	US-09-388-221-4	Sequence 12, Appli
41	9	0.7	1424	9	US-09-388-221-12	Sequence 2, Appli
42	9	0.7	1429	9	US-09-996-617-2	Sequence 2, Appli
43	9	0.7	1429	9	US-09-931-071-2	Sequence 5, Appli
44	9	0.7	1429	14	US-10-029-347-3	Sequence 26, Appli
45	9	0.7	1429	14	US-10-029-347-26	

## ALIGNMENTS

RESULT 1\nUS-10-124-498-6\n; Sequence 6, Application US/10124498\n; Publication No. US2003001983A1\nGENERAL INFORMATION:\n; APPLICANT: Bertin, John\n; APPLICANT: Wang, Weiye\n; APPLICANT: Blatcher, Maria\nTITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR\nFILE REFERENCE: 07334-367001\nCURRENT APPLICATION NUMBER: US/10/124,498\nPRIOR FILING DATE: 2002-04-17\nPRIOR APPLICATION NUMBER: 10/066,521\nPRIOR FILING DATE: 2002-01-31\nPRIOR APPLICATION NUMBER: 60/318,645\nPRIOR FILING DATE: 2001-09-10\nPRIOR APPLICATION NUMBER: 60/265,231\nPRIOR FILING DATE: 2001-01-31\nNUMBER OF SEQ ID NOS: 25\nSOFTWARE: FastSeq for Windows Version 4.0\nSEQ ID NO 6\nLENGTH: 1344\n; TYPE: PRT\n; ORGANISM: Homo sapiens\nUS-10-124-498-6\nQuery Match 100.0%; Score 1344; DB 14; Length 1344;\nBest Local Similarity 100.0%; Pred. No. 0;\nMatches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;\nCy 1 MEGDKSLTFSSYGLQWCLYEIDKEEPTFKELKKSSSTTCSIPQFIENAVVECTAL 60\nDb 1 MEGDKSLTFSSYGLQWCLYEIDKEEPTFKELKKSSSTTCSIPQFIENAVVECTAL 60

QY 61 LHHEYYGASIAMATSIISIFENNMLRTLSEKARDMKKISQAMEOGATAETEBEISQA 120  
DB 61 LHHEYYGASIAMATSIISIFENNMLRTLSEKARDMKKISQAMEOGATAETEBEISQA 120  
QY 121 MEQOGATAETEBEIGHGDDTMDYKSHVMTKFAEEDVRRSFENTAADPEMOTLAGAFDS 180  
DB 121 MEQOGATAETEBEIGHGDDTMDYKSHVMTKFAEEDVRRSFENTAADPEMOTLAGAFDS 180  
QY 181 DRMGFRPRTVVLHGSGIGKSALARRIVLCMAOGGLYQGMFSYVFPLVREMRKKESSV 240  
DB 181 DRMGFRPRTVVLHGSGIGKSALARRIVLCMAOGGLYQGMFSYVFPLVREMRKKESSV 240  
QY 241 TEFTSRWPDSQAPYTEIMSRPERLLFTIDGFDLGSVLANNDTKLCKDMAEKOPFTLIR 300  
DB 241 TEFTSRWPDSQAPYTEIMSRPERLLFTIDGFDLGSVLANNDTKLCKDMAEKOPFTLIR 300  
QY 301 SLRKVLLPESFLITVTRDVGTETKLSKVSPRYLLVRGISEQRHLLERIGEHOXT 360  
DB 301 SLRKVLLPESFLITVTRDVGTETKLSKVSPRYLLVRGISEQRHLLERIGEHOXT 360  
QY 361 QGLRAIMNNRELLDQCQVPAGSLICVALQLQDVVGESVAPPNQTLTGLHAAFFVHQLTP 420  
DB 361 QGLRAIMNNRELLDQCQVPAGSLICVALQLQDVVGESVAPPNQTLTGLHAAFFVHQLTP 420  
QY 421 RGVVRCCLNLEERVVLKFFCMAVEGVNNRKSVPFGDDLMOGLGESELRALFHMNILLP 480  
DB 421 RGVVRCCLNLEERVVLKFFCMAVEGVNNRKSVPFGDDLMOGLGESELRALFHMNILLP 480  
QY 481 DSHCEYVTFPHLSIQDFCAALYYVLEGLIEPALCPLYVEKTKSMELKQAGFIHSLW 540  
DB 481 DSHCEYVTFPHLSIQDFCAALYYVLEGLIEPALCPLYVEKTKSMELKQAGFIHSLW 540  
QY 541 MKRFLFGVSEBVRPRLVLLGCPPLGVKQKLLHWVLLGQOPAAATPGTTLAEHCLF 600  
DB 541 MKRFLFGVSEBVRPRLVLLGCPPLGVKQKLLHWVLLGQOPAAATPGTTLAEHCLF 600  
QY 601 ETODKEFVRLALNSRQEWLPIQNLIDLIASSFCLOHPYLRKIRVDYKGIFFRDESAA 660  
DB 601 ETODKEFVRLALNSRQEWLPIQNLIDLIASSFCLOHPYLRKIRVDYKGIFFRDESAA 660  
QY 661 CPVVDPLMRDKTLIEBQWEDFCSMLGTHPHLRQDLGSSILITERAKTLCAKLRHPTCKI 720  
DB 661 CPVVDPLMRDKTLIEBQWEDFCSMLGTHPHLRQDLGSSILITERAKTLCAKLRHPTCKI 720  
QY 721 QTLMRNAQITPGVOHLMRIWANNNRSLNIGTHLKEEDVRMACEALKPKCLBSLR 780  
DB 721 QTLMRNAQITPGVOHLMRIWANNNRSLNIGTHLKEEDVRMACEALKPKCLBSLR 780  
QY 781 LDCCGLTHACYLKIISQILTTSPSLKSLAGNKVTDGCMPLSDALRVSOQALQTLIED 840  
DB 781 LDCCGLTHACYLKIISQILTTSPSLKSLAGNKVTDGCMPLSDALRVSOQALQTLIED 840  
QY 841 CGITATGCGQSIALASALVSNRSLTHCLSNNSLGNBGNVLLCRSMRLPHCSQRLMNOCHL 900  
DB 841 CGITATGCGQSIALASALVSNRSLTHCLSNNSLGNBGNVLLCRSMRLPHCSQRLMNOCHL 900  
QY 901 DTAGGCFALALMGNSMLTHSLSNPVEDNGVLLCEVMRBPSCHODLELVKCHLAA 960  
DB 901 DTAGGCFALALMGNSMLTHSLSNPVEDNGVLLCEVMRBPSCHODLELVKCHLAA 960  
QY 961 CCESTLSCVTSRSLKSLDLTDNALGDGVAALCEGLKOKNSVLRGLKAKGSLTSDCE 1020  
DB 961 CCESTLSCVTSRSLKSLDLTDNALGDGVAALCEGLKOKNSVLRGLKAKGSLTSDCE 1020  
QY 1021 ALSTALSCNRHLTSLNLYVONNFSPPKMMKCSAFACPTSNLIIGLMTQYPPVOIRKLE 1080  
DB 1021 ALSTALSCNRHLTSLNLYVONNFSPPKMMKCSAFACPTSNLIIGLMTQYPPVOIRKLE 1080  
QY 1081 EVOLLKPRVVIDGSHSFDDEDRHKIGLTFRLPESRAPPCALLMWMNEOKKRVSLAGD 1140  
DB 1081 EVOLLKPRVVIDGSHSFDDEDRHKIGLTFRLPESRAPPCALLMWMNEOKKRVSLAGD 1140

QY 1141 FKSSTRPAKSLCLATANGESQVNDNEOSSPOPAAGTEHKKDLGVYSGAMSETAELE 1200  
DB 1141 FKSSTRPAKSLCLATANGESQVNDNEOSSPOPAAGTEHKKDLGVYSGAMSETAELE 1200  
QY 1201 GLGSNSADHDHGMAMSLGRELSRGICPTVLTMTTAVCPGHWERLGRGMCCLNSADHSG 1260  
DB 1201 GLGSNSADHDHGMAMSLGRELSRGICPTVLTMTTAVCPGHWERLGRGMCCLNSADHSG 1260  
QY 1261 VMSLSGAAGLEGLVNSADHDSGVAMSLGAAGLEGLVNSADHDSGVAMSLGAAGLEGLV 1320  
DB 1261 VMSLSGAAGLEGLVNSADHDSGVAMSLGAAGLEGLVNSADHDSGVAMSLGAAGLEGLV 1320  
QY 1321 SNSADHDSGVAMSLGAAGLEGLVNSADHDSGVAMSLGAAGLEGLV 1344  
DB 1321 SNSADHDSGVAMSLGAAGLEGLVNSADHDSGVAMSLGAAGLEGLV 1344

RESULT 2  
US-10-066-521-6  
; Sequence 6, Application US/1006521  
; Publication No. US20030027757A1  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; APPLICANT: Wang, Weiye  
; APPLICANT: Blacher, Maria  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR  
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF  
; FILE REFERENCE: 0734-334001  
; CURRENT APPLICATION NUMBER: US/10/066, 521  
; PRIOR FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: 60/318, 645  
; PRIOR FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/265, 231  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 1344  
; TYPE: PRN  
; ORGANISM: Homo sapiens  
US-10-066-521-6

Query Match 100.0%; Score 1344; DB 14; Length 1344;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGDKSLTFSSYGLQWCLYEIDKEEFOTFKELKKKSSBSTGCSIPQFIENANVECLAL 60  
DB 1 MEGDKSLTFSSYGLQWCLYEIDKEEFOTFKELKKKSSBSTGCSIPQFIENANVECLAL 60  
QY 61 LHHEYYGASIAMATSIISIFENNMLRTLSEKARDMKKISQAMEOGATAETEBEISQA 120  
DB 61 LHHEYYGASIAMATSIISIFENNMLRTLSEKARDMKKISQAMEOGATAETEBEISQA 120  
QY 121 MEQOGATAETEBEIGHGDDTMDYKSHVMTKFAEEDVRRSFENTAADPEMOTLAGAFDS 180  
DB 121 MEQOGATAETEBEIGHGDDTMDYKSHVMTKFAEEDVRRSFENTAADPEMOTLAGAFDS 180  
QY 181 DRMGFRPRTVVLHGSGIGKSALARRIVLCMAOGGLYQGMFSYVFPLVREMRKKESSV 240  
DB 181 DRMGFRPRTVVLHGSGIGKSALARRIVLCMAOGGLYQGMFSYVFPLVREMRKKESSV 240  
QY 241 TEFTSRWPDSQAPYTEIMSRPERLLFTIDGFDLGSVLANNDTKLCKDMAEKOPFTLIR 300  
DB 241 TEFTSRWPDSQAPYTEIMSRPERLLFTIDGFDLGSVLANNDTKLCKDMAEKOPFTLIR 300  
QY 301 SLRKVLLPESFLITVTRDVGTETKLSKVSPRYLLVRGISEQRHLLERIGEHOXT 360  
DB 301 SLRKVLLPESFLITVTRDVGTETKLSKVSPRYLLVRGISEQRHLLERIGEHOXT 360  
QY 361 QGLRAIMNNRELLDQCQVPAGSLICVALQLQDVVGESVAPPNQTLTGLHAAFFVHQLTP 420  
DB 361 QGLRAIMNNRELLDQCQVPAGSLICVALQLQDVVGESVAPPNQTLTGLHAAFFVHQLTP 420

```

QY 421 RGVNRCLNLEERVYLKRCRMAVEGVNRRKSVFDDDLINOGLEGSELRALFHHNLLP 480
DB 421 RGVNRCLNLEERVYLKRCRMAVEGVNRRKSVFDDDLINOGLEGSELRALFHHNLLP 480
QY 481 DSHCEEYTFPHLSLQDFCALYYVLLEGLIEPALCPLVEKTKSMELKQAGFIHSLM 540
DB 481 DSHCEEYTFPHLSLQDFCALYYVLLEGLIEPALCPLVEKTKSMELKQAGFIHSLM 540
QY 541 MKRPLFGLVSEVDYRPLEVLLGCPVPLGVKOKLLHNVSLLGQOPNATTEGDTLDAFHCLE 600
DB 541 MKRPLFGLVSEVDYRPLEVLLGCPVPLGVKOKLLHNVSLLGQOPNATTEGDTLDAFHCLE 600
QY 601 ETQDEPFRALANSGOEVLPIPNQDLIASSFCLOHCYLAIRVDVVGIFPRDESABA 660
DB 601 ETQDEPFRALANSGOEVLPIPNQDLIASSFCLOHCYLAIRVDVVGIFPRDESABA 660
QY 661 CPVPLMWRDKTLIEQWEDFCSMLGTHPHLRQLDSSILTERAMKTLCAKLHPPTCKI 720
DB 661 CPVPLMWRDKTLIEQWEDFCSMLGTHPHLRQLDSSILTERAMKTLCAKLHPPTCKI 720
QY 721 QTLMRNMQITPVOHLMRIYVANNRLSLNLGGTHLKEEDVRMACEALKHPKCLLESIR 780
DB 721 QTLMRNMQITPVOHLMRIYVANNRLSLNLGGTHLKEEDVRMACEALKHPKCLLESIR 780
QY 781 LDCCGLTHACYLIKISQILITTSPLSKSLAGKNTVDQVMPISDALRVSGCALQKLLIED 840
DB 781 LDCCGLTHACYLIKISQILITTSPLSKSLAGKNTVDQVMPISDALRVSGCALQKLLIED 840
QY 841 CGITANGCQSLASALVSNRSLTHLCLSNNSLGNBGNVLCRSMRLPHCSLQRLMNAQCHL 900
DB 841 CGITANGCQSLASALVSNRSLTHLCLSNNSLGNBGNVLCRSMRLPHCSLQRLMNAQCHL 900
QY 901 DTAGCGFLALAMGNSWLTHLSLSPNPVEDNGVKLLCEVWRPESCHLQDLIELVKCHLTAA 960
DB 901 DTAGCGFLALAMGNSWLTHLSLSPNPVEDNGVKLLCEVWRPESCHLQDLIELVKCHLTAA 960
QY 961 CCEBSLSCVTSRSHLKSJDLTDNALGDGVVALCEGLKOKNSVLTRLGKACGLTSDCE 1020
DB 961 CCEBSLSCVTSRSHLKSJDLTDNALGDGVVALCEGLKOKNSVLTRLGKACGLTSDCE 1020
QY 1021 ALSLASCNRHLTSLNLYONNPSPKGMMLCSAPACPTSNLQITGLMKQYVQIRKYLE 1080
DB 1021 ALSLASCNRHLTSLNLYONNPSPKGMMLCSAPACPTSNLQITGLMKQYVQIRKYLE 1080
QY 1081 EYQOLKPRVVIDSGWHSFEDDRHKIGLTFRLPESRAMPCALLMGNNPEQKRVSLLAGD 1140
DB 1081 EYQOLKPRVVIDSGWHSFEDDRHKIGLTFRLPESRAMPCALLMGNNPEQKRVSLLAGD 1140
QY 1141 FKSSSTRFPAKSLCLATANGESQKRVNVBOSSPOPMAGTEBKODKMLSVGSGAMSETAELE 1200
DB 1141 FKSSSTRFPAKSLCLATANGESQKRVNVBOSSPOPMAGTEBKODKMLSVGSGAMSETAELE 1200
QY 1201 GIGNSADHDGGMAMSLGRELSRGLCPTVMTTAAVCPGHWERLGSRCWCLNSADDSHG 1260
DB 1201 GIGNSADHDGGMAMSLGRELSRGLCPTVMTTAAVCPGHWERLGSRCWCLNSADDSHG 1260
QY 1261 VSMISGAAGLEGLVSNADHSGVAMSLGAAGLEGLVSNADHSGVAMSLGAAGLEGLV 1320
DB 1261 VSMISGAAGLEGLVSNADHSGVAMSLGAAGLEGLVSNADHSGVAMSLGAAGLEGLV 1320
QY 1321 SNSADHSGVAMSLGAAGLEGLV 1344
DB 1321 SNSADHSGVAMSLGAAGLEGLV 1344

```

```

RESULT 3
US-10-216-645-2
; Sequence 2, Application US/10216645
; Publication No. US20030125282A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; APPLICANT: LESSLI, MONIKA

```

```

; APPLICANT: PETERS-KOTTIG, MICHAEL
; APPLICANT: BECKMANN, GEORG
; TITLE OF INVENTION: HUMAN MATER PROTEINS
; FILE REFERENCE: SCH-1910
; CURRENT APPLICATION NUMBER: US/10/216, 645
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-216-645-2

Query Match      74.9%; Score 1006; DB 14; Length 1162;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1006; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 ISQAMEQCATPAETEEQEOISQAMEQCATPAETEEQHGDDTDYKSHWTKFAEEDV 157
DB 152 ISQAMEQCATPAETEEQEOISQAMEQCATPAETEEQHGDDTDYKSHWTKFAEEDV 211
QY 158 RRSFENTADMPENQTLGAFDSRWGFRPTVVLHGSGIGKSALARIVLCMAQGGLY 217
DB 212 RRSFENTADMPENQTLGAFDSRWGFRPTVVLHGSGIGKSALARIVLCMAQGGLY 271
QY 218 QGMFSYVFFLVRERQKKESSVTEFISREMPDSQAPTEIMSRPERLLIIDGFDDIGS 277
DB 272 QGMFSYVFFLVRERQKKESSVTEFISREMPDSQAPTEIMSRPERLLIIDGFDDIGS 331
QY 278 VLANDTKLCKDMAEKOPFTLLIRSLRKVLLPESFLIVTVADVTEKLSKEVSPRYLLV 337
DB 332 VLANDTKLCKDMAEKOPFTLLIRSLRKVLLPESFLIVTVADVTEKLSKEVSPRYLLV 391
QY 338 RGISGEQRIHLLERIGEGHQKQGLRAIMNRRELLDOCVPAVGSILCVALLQODVVG 397
DB 392 RGISGEQRIHLLERIGEGHQKQGLRAIMNRRELLDOCVPAVGSILCVALLQODVVG 451
QY 398 SVAPENQTLTGILAAFAVPHQTPRGVVRCLNLEERVYLKRCRMAVEGVNRRKSVFDDG 457
DB 452 SVAPENQTLTGILAAFAVPHQTPRGVVRCLNLEERVYLKRCRMAVEGVNRRKSVFDDG 511
QY 458 DLMVQGLGESELRALFHHNLLPDSHCEYTFPHLSLQDFCALYYVLLEGLIEPALCP 517
DB 512 DLMVQGLGESELRALFHHNLLPDSHCEYTFPHLSLQDFCALYYVLLEGLIEPALCP 571
QY 518 LYVEKTKRSMELKQAGFIHSLMKRFLFGLVSEVDYRPLEVLLGCPVPLGVKOKLLHNV 577
DB 572 LYVEKTKRSMELKQAGFIHSLMKRFLFGLVSEVDYRPLEVLLGCPVPLGVKOKLLHNV 631
QY 578 SLLGQOPNATTEGDTLDAFHCLEFETQDEPFRALANSGOEVLPIPNQDLIASSFCLOH 637
DB 632 SLLGQOPNATTEGDTLDAFHCLEFETQDEPFRALANSGOEVLPIPNQDLIASSFCLOH 691
QY 638 CPYLKIRIVDKGIFPRDESABACPVVPLMWRDKTLIEQWEDFCSMLGTHPHLRQLDGL 697
DB 692 CPYLKIRIVDKGIFPRDESABACPVVPLMWRDKTLIEQWEDFCSMLGTHPHLRQLDGL 751
QY 698 SSILTERAMKTLCAKLHPPTCKIOTLMFRNAQITPVOHLMRIYVANNRLSLNLGGTHL 757
DB 752 SSILTERAMKTLCAKLHPPTCKIOTLMFRNAQITPVOHLMRIYVANNRLSLNLGGTHL 811
QY 758 KEBDYRMACEALKHPKCLLESIRLDCCGLTHACYLIKISQILITTSPLSKSLAGKNTVDQ 817
DB 812 KEBDYRMACEALKHPKCLLESIRLDCCGLTHACYLIKISQILITTSPLSKSLAGKNTVDQ 871
QY 818 GVMPLSDALRVSGCALQKLLIEDCGITATGQSLASALVSNRSLTHLCLSNNSLGNBGN 877
DB 872 GVMPLSDALRVSGCALQKLLIEDCGITATGQSLASALVSNRSLTHLCLSNNSLGNBGN 931
QY 878 LILCRSMRLPHCSLQRLMNAQCHLDTAGCGFLALAMGNSWLTHLSLSPNPVEDNGVKLLC 937

```



Db 992 LLCRSMRLPHCSLQRLMLNQCGLDTAGCGFLALALMGNSWLTHLSLSMNPVEDNGVLLC 991  
Qy 938 EVMREPSCHLQDLVYKCHLTAACCESSCVTSRBRHLKSLDTLNALGDGVALCEGL 997  
Db 992 EVMREPSCHLQDLVYKCHLTAACCESSCVTSRBRHLKSLDTLNALGDGVALCEGL 1051  
Qy 998 KOKNSVLRGLGKACGLTSDCCCEALSLASCNRHLTSLNLYONNPSPKGMKLCSAFACP 1057  
Db 1052 KOKNSVLRGLGKACGLTSDCCCEALSLASCNRHLTSLNLYONNPSPKGMKLCSAFACP 1111  
Qy 1058 TSNLIIGLWKQYFVQIRKLLBEVOLLKPRVVIDGSMHSEDEDR 1103  
Db 1112 TSNLIIGLWKQYFVQIRKLLBEVOLLKPRVVIDGSMHSEDEDR 1157

RESULT 4  
US-10-416-642-1  
; Sequence 1, Application US/10416642  
; Publication No. US20040043452A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: RAMKUMAR, Jayajakmi  
; APPLICANT: ARVIZU, Chandra  
; TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0842 PCT  
; CURRENT APPLICATION NUMBER: US/10/416,642  
; CURRENT FILING DATE: 2003-05-13  
; PRIOR APPLICATION NUMBER: 60/249,407  
; PRIOR FILING DATE: 2000-11-15  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PERL Program  
; SEQ ID NO 1  
; LENGTH: 1162  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040043452A1 7474830CD1  
US-10-416-642-1

Query Match 74.9%; Score 1006; DB 15; Length 1162;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1006; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 98 ISQAMEOGATTAETEOEISQAMEOGATTAETEOHGTWYKSHWTKFAEEEDV 157  
Db 152 ISQAMEOGATTAETEOEISQAMEOGATTAETEOHGTWYKSHWTKFAEEEDV 211  
Qy 158 RRSFNTAADWPEMOTLAGAFSDRMGFRPTVVLHGSGIGKSLARIVLCMAOGGLY 217  
Db 212 RRSFNTAADWPEMOTLAGAFSDRMGFRPTVVLHGSGIGKSLARIVLCMAOGGLY 271  
Qy 218 QGMFYSVFLLPVREHQRKKESSVTEFISREWDSQAPTEIMSRPERLLFTIDGFDIGS 277  
Db 272 QGMFYSVFLLPVREHQRKKESSVTEFISREWDSQAPTEIMSRPERLLFTIDGFDIGS 331  
Qy 278 VLNNTTKLCKMAEKOPPTLLRSLLRKVLLPESFLIVTVRVGYTEKLSKEVSVPRYLLV 337  
Db 332 VLNNTTKLCKMAEKOPPTLLRSLLRKVLLPESFLIVTVRVGYTEKLSKEVSVPRYLLV 391  
Qy 338 RGISGEORIHLLERIGIEHOKTQGLRAIMNNRELLDCCQVPVAVSLLICVALQLDVGGE 397  
Db 392 RGISGEORIHLLERIGIEHOKTQGLRAIMNNRELLDCCQVPVAVSLLICVALQLDVGGE 451  
Qy 398 SVAPFNQTLTGHAAPVHFQLTLPBGVVRCLNLEBRVYLKEPCRMAVEGWNRRKSVDPGD 457  
Db 452 SVAPFNQTLTGHAAPVHFQLTLPBGVVRCLNLEBRVYLKEPCRMAVEGWNRRKSVDPGD 511  
Qy 458 DLMVGLSESELRALFHNHLLPDSHCERYTFHFLLSLQDPCALYYTLLEGITPALCP 517  
Db 512 DLMVGLSESELRALFHNHLLPDSHCERYTFHFLLSLQDPCALYYTLLEGITPALCP 571

Qy 518 LYVEKTRSMELKQAGFHISHLMKREPLFGLVSEDEVRRPLEVLLGCPVPLGVKOKLHMV 577  
Db 572 LYVEKTRSMELKQAGFHISHLMKREPLFGLVSEDEVRRPLEVLLGCPVPLGVKOKLHMV 631  
Qy 578 SLIQOQPNATTPGDTLDAFHCLEFETOKEFVRALNSFOEYVLPINONLDLIASSFCLOH 637  
Db 632 SLIQOQPNATTPGDTLDAFHCLEFETOKEFVRALNSFOEYVLPINONLDLIASSFCLOH 691  
Qy 638 CPYLRKTRVDYKGIFFPDESSEAECPPVPLMWBDTLLEBOWEDCSMLGTHPHLRQDLG 697  
Db 692 CPYLRKTRVDYKGIFFPDESSEAECPPVPLMWBDTLLEBOWEDCSMLGTHPHLRQDLG 751  
Qy 698 SSILTERAMKTLCAKLHPCTCKIOTLMFRNAQITPGVQHLMRIYMANRNLRSNLGGTHL 757  
Db 752 SSILTERAMKTLCAKLHPCTCKIOTLMFRNAQITPGVQHLMRIYMANRNLRSNLGGTHL 811  
Qy 758 KEBDVNACELAKPKCLBSLRDCCGLTHACTYKLSQILITSPSKSLAGNKYTDQ 817  
Db 812 KEBDVNACELAKPKCLBSLRDCCGLTHACTYKLSQILITSPSKSLAGNKYTDQ 871  
Qy 818 GVMPLSDALRYSCALQKLLIEDCGITATGCGSLASALVSNRSLTHCLSNNSLGNBGN 877  
Db 872 GVMPLSDALRYSCALQKLLIEDCGITATGCGSLASALVSNRSLTHCLSNNSLGNBGN 931  
Qy 878 LLCRSMRLPHCSLQRLMLNQCGLDTAGCGFLALALMGNSWLTHLSLSMNPVEDNGVLLC 937  
Db 932 LLCRSMRLPHCSLQRLMLNQCGLDTAGCGFLALALMGNSWLTHLSLSMNPVEDNGVLLC 991  
Qy 938 EVMREPSCHLQDLVYKCHLTAACCESSCVTSRBRHLKSLDTLNALGDGVALCEGL 997  
Db 992 EVMREPSCHLQDLVYKCHLTAACCESSCVTSRBRHLKSLDTLNALGDGVALCEGL 1051  
Qy 998 KOKNSVLRGLGKACGLTSDCCCEALSLASCNRHLTSLNLYONNPSPKGMKLCSAFACP 1057  
Db 1052 KOKNSVLRGLGKACGLTSDCCCEALSLASCNRHLTSLNLYONNPSPKGMKLCSAFACP 1111  
Qy 1058 TSNLIIGLWKQYFVQIRKLLBEVOLLKPRVVIDGSMHSEDEDR 1103  
Db 1112 TSNLIIGLWKQYFVQIRKLLBEVOLLKPRVVIDGSMHSEDEDR 1157

RESULT 5  
US-10-216-645-4  
; Sequence 4, Application US/10216645  
; Publication No. US20030125282A1  
; GENERAL INFORMATION:  
; APPLICANT: WEISS, BERTRAM  
; APPLICANT: LESSL, MONIKA  
; APPLICANT: PETERS-KOTTIG, MICHAEL  
; APPLICANT: BECKMANN, GEORG  
; TITLE OF INVENTION: HUMAN MATER PROTEINS  
; FILE REFERENCE: SCH-1910  
; CURRENT APPLICATION NUMBER: US/10/216,645  
; CURRENT FILING DATE: 2003-01-21  
; PRIOR APPLICATION NUMBER: DE 101 39 874.3 41  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 1143  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-216-645-4

Query Match 73.5%; Score 988; DB 14; Length 1143;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 988; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 EISQAMEOGATTAETEOHGTWYKSHWTKFAEEEDVRRSFNTAADWPEMOTLA 175  
Db 151 EISQAMEOGATTAETEOHGTWYKSHWTKFAEEEDVRRSFNTAADWPEMOTLA 210  
Qy 176 GAFSDRMGFRPTVVLHGSGIGKSLARIVLCMAOGGLYQGMFYSVFLLPVREHQRK 235

```

Db 211 GAFSDRNGFRRTVVLHGKSGIGKSLARRIVLCMAQGGYQGMFSYVFLPVREMRK 270
Qy 236 KESSTERTISREWPDSQAPVTEINRPERELTITGFPDLSVANNDTCLCDMAEKOP 295
Db 271 KESSTERTISREWPDSQAPVTEINRPERELTITGFPDLSVANNDTCLCDMAEKOP 330
Qy 296 FTLLSLRKVLLPSSFLIVTVRDVTEKSKSEVSPRYLVRGISEGRHLLERGG 355
Db 331 FTLLSLRKVLLPSSFLIVTVRDVTEKSKSEVSPRYLVRGISEGRHLLERGG 390
Qy 356 EHOKTQGLRAIMNNELLDOCVPAVGSILCVALQLODVGSVAPENOTLTLHAAYF 415
Db 391 EHOKTQGLRAIMNNELLDOCVPAVGSILCVALQLODVGSVAPENOTLTLHAAYF 450
Qy 416 HOLTRGVRRCLNLEERVVTKRCPMAVEGWNRKSVFDDGDLWVQGLSESELALFPM 475
Db 451 HOLTRGVRRCLNLEERVVTKRCPMAVEGWNRKSVFDDGDLWVQGLSESELALFPM 510
Qy 476 NILPDSHCEEYTFPHSLDPCALYYVLEGLTEIPRLCPLVYKTKRSMELKQAGH 535
Db 511 NILPDSHCEEYTFPHSLDPCALYYVLEGLTEIPRLCPLVYKTKRSMELKQAGH 570
Qy 536 IHSLMKRFGLVSEVDVRRPLEVLLGCEVPLGVKQKLLHWVSLGQOPNATTPGDTLDA 595
Db 571 IHSLMKRFGLVSEVDVRRPLEVLLGCEVPLGVKQKLLHWVSLGQOPNATTPGDTLDA 630
Qy 596 FHCLFETDKEFVRILNLSFOEWPMLPINONDLIASSFCLQHCPLYRKIRVDVKIIPRD 655
Db 631 FHCLFETDKEFVRILNLSFOEWPMLPINONDLIASSFCLQHCPLYRKIRVDVKIIPRD 690
Qy 656 ESABCPVPLMMRKTLIEQWEDPCSLGTHPHLRQDLGSSILTERAMTKLCAKLRH 715
Db 691 ESABCPVPLMMRKTLIEQWEDPCSLGTHPHLRQDLGSSILTERAMTKLCAKLRH 750
Qy 716 PTCKIQTLMFRNAQITPGVQHLRIYMANRNLRSNLGTHKEEDVMAACEALGHPKCL 775
Db 751 PTCKIQTLMFRNAQITPGVQHLRIYMANRNLRSNLGTHKEEDVMAACEALGHPKCL 810
Qy 776 LESRLDCCGLTHACYLKISQILTTSPSLKSLAGNVTGQVPLSDALRVSQCALQK 835
Db 811 LESRLDCCGLTHACYLKISQILTTSPSLKSLAGNVTGQVPLSDALRVSQCALQK 870
Qy 836 LILBECGTTATGCCGSLASALVSNRSLTHCLSNNSLGNBVNLLCRSRLPHCSLQRLML 895
Db 871 LILBECGTTATGCCGSLASALVSNRSLTHCLSNNSLGNBVNLLCRSRLPHCSLQRLML 930
Qy 896 NOCHDPTAGCGFLATLALMGNSWLTHLSLMPNVEDNGVYLLCEVNRBPSCHLQDELYKC 955
Db 931 NOCHDPTAGCGFLATLALMGNSWLTHLSLMPNVEDNGVYLLCEVNRBPSCHLQDELYKC 990
Qy 956 HLTAAACESLSVCISRSRHLKSLDLTDNALGDGVAALCEGLKQKNSVLTBLGLKACGLT 1015
Db 991 HLTAAACESLSVCISRSRHLKSLDLTDNALGDGVAALCEGLKQKNSVLTBLGLKACGLT 1050
Qy 1016 SDCCALSLAISCNHLLTSLNVQNNFSPKGMKCSAPACTSNLQITIGLKKQYPIQI 1075
Db 1051 SDCCALSLAISCNHLLTSLNVQNNFSPKGMKCSAPACTSNLQITIGLKKQYPIQI 1110
Qy 1076 RKLLEEVOLLKPRVYIDGSMHSFDEDDR 1103
Db 1111 RKLLEEVOLLKPRVYIDGSMHSFDEDDR 1138

```

RESULT 6  
US-10-092-900A-348  
Sequence 348. Application US/10092900A  
Publication No. US2004004382A1

GENERAL INFORMATION:  
APPLICANT: Padigar, Muralidhara  
APPLICANT: Sanyal, Kimberly A.  
APPLICANT: Shenoy, Suresh G.  
APPLICANT: Taupier Jr., Raymond J.

```

APPLICANT: Pena, Carol E.A.
APPLICANT: Li, Li
APPLICANT: Zernusen, Bryan D.
APPLICANT: Gueev, Vladimir Y.
APPLICANT: Ji, Weizhen
APPLICANT: Gorman, Linda
APPLICANT: Miller, Charles E.
APPLICANT: Kekuda, Ramesh
APPLICANT: Paturajan, Meera
APPLICANT: Gangoli, Bha A.
APPLICANT: Verne, Corine A.M.
APPLICANT: Guo, Xiaojia Sashia
APPLICANT: Tchenev, Velizar T.
APPLICANT: Fernandes, Elma R.
APPLICANT: Casman, Stacie J.
APPLICANT: Malyankar, Uriel M.
APPLICANT: Gerlach, Valerie
APPLICANT: Liu, Yi
APPLICANT: Anderson, David W.
APPLICANT: Spaderna, Steven K.
APPLICANT: Catterton, Elina
APPLICANT: Lette, Mario W.
APPLICANT: Zhong, Haihong
APPLICANT: Alsobrook, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
FILE OF INVENTION: No. US2004004382A1 Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-290C
CURRENT APPLICATION NUMBER: US/10/092,900A
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: USN 60/274,322
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USN 60/283,675
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: USN 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: USN 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USN 60/274,191
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USN 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: USN 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: USN 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: USN 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: USN 60/287,424
PRIOR FILING DATE: 2001-04-30
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 768
SEQ ID NO 348
LENGTH: 1033
TYPE: PRT
ORGANISM: Homo sapiens
US-10-092-900A-348

```

```

Query Match 72.0%; Score 968; DB 15; Length 1033;
Query Local Similarity 100.0%; Pred. No. 0;
Matches 968; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 136 HGGDTWDYKSHVMTKFAEEDVRRSPENTADWPEMOTLAGAPSDRNGFRRTVVLHGK 195
Db 61 HGGDTWDYKSHVMTKFAEEDVRRSPENTADWPEMOTLAGAPSDRNGFRRTVVLHGK 120
Qy 196 SGIGKSALARRIVLCMAQGGYQGMFSYVFLPVREMRKESSTERTISREWPDSQAPV 255
Db 121 SGIGKSALARRIVLCMAQGGYQGMFSYVFLPVREMRKESSTERTISREWPDSQAPV 180
Qy 256 TEISRPERLLFIIDGFPDLSVANNDTKCLCDMAEKOPPTLIRSLRKVLLPESFLIV 315

```

181 TEIMSRPERLLFIIDGFDLGSVLNNDRKLCOMAEKOPFETLLIRSLRKVLLPESFLIV 240  
QY 316 TWRDGTSEKTKSEVVSPLYLVKRGISGEORHLLERIGEHOKTOGRALMNNRELLDQ 375  
Db 241 TWRDGTSEKTKSEVVSPLYLVKRGISGEORHLLERIGEHOKTOGRALMNNRELLDQ 300  
QY 376 CQVPVAGSLICVALQLODVGVESVAPFNQTLTGHAFAVFHQLTRGVVRCCLNEERV 435  
Db 301 CQVPVAGSLICVALQLODVGVESVAPFNQTLTGHAFAVFHQLTRGVVRCCLNEERV 360  
QY 436 LKRFRCMAVEGVWNRKSVFDDGLMVOGLGESELPALFHMNLLPDSHCEBYTTFHSL 495  
Db 361 LKRFRCMAVEGVWNRKSVFDDGLMVOGLGESELPALFHMNLLPDSHCEBYTTFHSL 420  
QY 496 ODFCALYYVLEGLIEPACPLVYEKTKRSMELKQAGPHIHSIMMKKFLFGLVSEDRR 555  
Db 421 ODFCALYYVLEGLIEPACPLVYEKTKRSMELKQAGPHIHSIMMKKFLFGLVSEDRR 480  
QY 556 PLEVLLGCPVPLGVKQKLLHWVSLGQOPNATTPGDTLDAFHCLFETODKEFVRALNSF 615  
Db 481 PLEVLLGCPVPLGVKQKLLHWVSLGQOPNATTPGDTLDAFHCLFETODKEFVRALNSF 540  
QY 616 QEWPLPINQNLDLIASSFCLQHCPLYRKIRVDVKGI PRDESSEAACPVPIMMRDXTLIE 675  
Db 541 QEWPLPINQNLDLIASSFCLQHCPLYRKIRVDVKGI PRDESSEAACPVPIMMRDXTLIE 600  
QY 676 EOMEPFCMLGTHPHLRQDLGSSILTERAMKTLCAKLRHPTCKIOTLMFNAQITTPVQ 735  
Db 601 EOMEPFCMLGTHPHLRQDLGSSILTERAMKTLCAKLRHPTCKIOTLMFNAQITTPVQ 660  
QY 736 HLMRIYMANRNLRSINLGTGTHLEEDVMAACEALAKHPKCLLESRLDDCCGLTHACYLKIS 795  
Db 661 HLMRIYMANRNLRSINLGTGTHLEEDVMAACEALAKHPKCLLESRLDDCCGLTHACYLKIS 720  
QY 796 QILTTSPSLKISLTAQNKVTDGCVMPPLDALRVSCALOKLILEDCGITTATGCCSLASAL 855  
Db 721 QILTTSPSLKISLTAQNKVTDGCVMPPLDALRVSCALOKLILEDCGITTATGCCSLASAL 780  
QY 856 VSNRSLTJLCLSNNSLGNBGNVLLCRSMRLPHCSIORLMLNOCHDPTAGCGFLAALMGN 915  
Db 781 VSNRSLTJLCLSNNSLGNBGNVLLCRSMRLPHCSIORLMLNOCHDPTAGCGFLAALMGN 840  
QY 916 SWLTLTSLSMNVEEDNGVYLLCEVWRBPSCHLQDLELVKCHLTAACCSLSVCISRSRHL 975  
Db 841 SWLTLTSLSMNVEEDNGVYLLCEVWRBPSCHLQDLELVKCHLTAACCSLSVCISRSRHL 900  
QY 976 KSLDLTDNALGPGVYALCEGLKQKNSVLTIRGLAKCGITSDCCALSLASCNHRLTSL 1035  
Db 901 KSLDLTDNALGPGVYALCEGLKQKNSVLTIRGLAKCGITSDCCALSLASCNHRLTSL 960  
QY 1036 NLVONNPSFKGMKLCAPACPTSNLQIIGLMMKQYPOAIRKLEEVOLLRVVIDGSM 1095  
Db 961 NLVONNPSFKGMKLCAPACPTSNLQIIGLMMKQYPOAIRKLEEVOLLRVVIDGSM 1020  
QY 1096 HSFDEDDR 1103  
Db 1021 HSFDEDDR 1028

RESULT 7  
US-10-399-443-24  
; Sequence 24, Application US/10399443  
; Publication No. US20040028669A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America, as Represented by the  
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
; APPLICANT: Health  
; APPLICANT: Nelson, Lawrence M.  
; APPLICANT: Tong, Zhi-Bin  
; APPLICANT: Nelson, Lawrence  
; APPLICANT: Zhi-Bin, Tong  
; TITLE OF INVENTION: Human Gene Critical to Fertility  
; FILE REFERENCE: 4239-64785

; CURRENT APPLICATION NUMBER: US/10/399, 443  
; CURRENT FILING DATE: 2003-04-16  
; PRIOR APPLICATION NUMBER: 60/241, 510  
; PRIOR FILING DATE: 2000-10-18  
; PRIOR APPLICATION NUMBER: PCT/US01/10981  
; PRIOR FILING DATE: 2001-04-04  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 1200  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-399-443-24

Query Match 44.8%; Score 602; DB 15; Length 1200;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1002; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 98 ISQAMEOGATAATTEBOEISQAMEOGATTAETEOGGGDTWDYKSHWTKFAEBEDV 157  
Db 190 ISQAMEOGATAATTEBOEISQAMEOGATTAETEOGGGDTWDYKSHWTKFAEBEDV 249  
QY 158 RRSFENTAADWPEMOTLAGAFSDRWGFRPRTVVILHKSIGKSALARIVLCWAQGLY 217  
Db 250 RRSFENTAADWPEMOTLAGAFSDRWGFRPRTVVILHKSIGKSALARIVLCWAQGLY 309  
QY 218 QGMSSYFPLPYRMOKKSSVTEFTSRMPDQAVTEIMSRPERLLFIIDGFDLGS 277  
Db 310 QGMSSYFPLPYRMOKKSSVTEFTSRMPDQAVTEIMSRPERLLFIIDGFDLGS 369  
QY 278 VLANDTKLCKDMAEKOPFETLLIRSLRKVLLPESFLIVTVADVTEKTKSEVVSPLYLV 337  
Db 370 VLANDTKLCKDMAEKOPFETLLIRSLRKVLLPESFLIVTVADVTEKTKSEVVSPLYLV 429  
QY 338 RGISEGRIHLLERIGEHOKTOGRALMNNRELLDQCVPAVAGSLICVALQLODVVGE 397  
Db 430 RGISEGRIHLLERIGEHOKTOGRALMNNRELLDQCVPAVAGSLICVALQLODVVGE 489  
QY 398 SVAPFNQTLTGHAFAVFHQLTRGVVRCCLNEERVVLRFCMAVEGVWNRKSVFDDG 457  
Db 490 SVAPFNQTLTGHAFAVFHQLTRGVVRCCLNEERVVLRFCMAVEGVWNRKSVFDDG 549  
QY 458 DLMQGLGESELRALFHMNLLPDSHCEBYTTFHSLQDFCALYYVLEGLIEPACPL 517  
Db 550 DLMQGLGESELRALFHMNLLPDSHCEBYTTFHSLQDFCALYYVLEGLIEPACPL 609  
QY 518 LYVEKTKRSMELKQAGPHIHSIMMKRFLFGLVSEDRRPLEVLLGCPVPLGVKQKLLHW 577  
Db 610 LYVEKTKRSMELKQAGPHIHSIMMKRFLFGLVSEDRRPLEVLLGCPVPLGVKQKLLHW 669  
QY 578 SILGQOPNATTPGDTLDAFHCLFETODKEFVRALNSFOEYWLPIQNLDLIASSFCLQ 637  
Db 670 SILGQOPNATTPGDTLDAFHCLFETODKEFVRALNSFOEYWLPIQNLDLIASSFCLQ 729  
QY 638 CPYLARKIRVDVKGI PRDESSEAACPVPIMMRDXTLIEBQWEDFCMLGTHPHLRQDLG 697  
Db 730 CPYLARKIRVDVKGI PRDESSEAACPVPIMMRDXTLIEBQWEDFCMLGTHPHLRQDLG 789  
QY 698 SSILTERAMKTLCAKLRHPTCKIOTLMFNAQITTPVQHLMRIYMANRNLRSINLGTGTH 757  
Db 790 SSILTERAMKTLCAKLRHPTCKIOTLMFNAQITTPVQHLMRIYMANRNLRSINLGTGTH 849  
QY 758 KEEBVRMACEALAKHPKCLLESRLDDCCGLTHACYLKISQILTTSPSLKISLACNKVTDQ 817  
Db 850 KEEBVRMACEALAKHPKCLLESRLDDCCGLTHACYLKISQILTTSPSLKISLACNKVTDQ 909  
QY 818 GMPPLDALRVSCALOKLILEDCGITTATGCCSLASLVNSRSLTJLCLSNNSLGNBGNV 877  
Db 910 GMPPLDALRVSCALOKLILEDCGITTATGCCSLASLVNSRSLTJLCLSNNSLGNBGNV 969  
QY 878 LKCRSMRLPHCSIORLMLNOCHDPTAGCGFLAALMGNSWLTLTSLSMNVEEDNGVYLLC 937  
Db 970 LKCRSMRLPHCSIORLMLNOCHDPTAGCGFLAALMGNSWLTLTSLSMNVEEDNGVYLLC 1029

Qy	998	EVMRPSCHLODELVKCHLTACCESSJCSVSRSHKLSLJLTNALGDGVAALCBEL	999
Db	1030	EVMRPSCHLODELVKCHLTACCESSJCSVSRSHKLSLJLTNALGDGVAALCBEL	1088
Qy	998	KOKNSVTLRLGKACGLTSDCCALSLALSCNRHJLSJLTVONNPSPKMMKLCSAFACP	105
Db	1090	KOKNSVTLRLGKACGLTSDCCALSLALSCNRHJLSJLTVONNPSPKMMKLCSAFACP	1144
Qy	1058	TSNLOITIGMKQYPVQIRKLEEVQLKPRVYVIGSWHSFPEDDR	1103
Db	1150	TSNLOITIGMKQYPVQIRKLEEVQLKPRVYVIGSWHSFPEDDR	1195
RESULT 8			
US-10-677-943-24			
Sequence 24, Application US/10677943			
Publication No. US20040072297A1			
GENERAL INFORMATION:			
APPLICANT: The Government of the United States of America as			
APPLICANT: represented by the Secretary of the Department of Health and			
APPLICANT: Human Services			
APPLICANT: Nelson, Lawrence			
APPLICANT: Tong, Zhi-Bin			
TITLE OF INVENTION: Human Gene Critical to Fertility			
FILE REFERENCE: 4239-64790			
CURRENT APPLICATION NUMBER: US/10/677, 943			
PRIOR FILING DATE: 2003-10-01			
PRIOR APPLICATION NUMBER: 60/241,510			
PRIOR FILING DATE: 2000-10-18			
PRIOR APPLICATION NUMBER: PCT/US02/09776			
PRIOR FILING DATE: 2002-03-29			
PRIOR APPLICATION NUMBER: PCT/US01/10981			
PRIOR FILING DATE: 2001-04-04			
NUMBER OF SEQ ID NOS: 42			
SOFTWARE: PatentIn version 3.2			
SEQ ID NO: 24			
LENGTH: 1200			
TYPE: PRT			
ORGANISM: Homo sapiens			
US-10-677-943-24			
Query Match			
Best Local Similarity 99.6%; Score 602; DB 15; Length 1200;			
Matches 1002; Conservative 0; Mismatches 4; Indels 0; Gaps 0			
Qy	98	ISQAMEQCATTAETEEBOEISQAMEQSGATTAETEEQHGSDJTWYKSHVMTKFAEBDV	157
Db	190	ISQAMEQCATTAETEEBOEISQAMEQSGATTAETEEQHGSDJTWYKSHVMTKFAEBDV	249
Qy	158	RRSFENTADMPENOTLGAAPDSMDWGFRRPVVLVHGSGIGKSLARIVLCMAOGGLY	217
Db	250	RRSFENTADMPENOTLGAAPDSMDWGFRRPVVLVHGSGIGKSLARIVLCMAOGGLY	309
Qy	218	QGMFSYVFLPVREMKRKSSVTEFISREWPDSQAPYTEIMSRPERLLFIIDGFDLGS	277
Db	310	QGMFSYVFLPVREMKRKSSVTEFISREWPDSQAPYTEIMSRPERLLFIIDGFDLGS	369
Qy	278	VLANDTKLCCKMAEQAPFTLIRSLRLKRVLLPESLIVTVRDVGTCKLSEVYSPRYLLV	337
Db	370	VLANDTKLCCKMAEQAPFTLIRSLRLKRVLLPESLIVTVRDVGTCKLSEVYSPRYLLV	429
Qy	338	RGISGEQRIHLLERIGIEGHOKTQGLRAIMNNRELLDDCOQVAVSGLICVALQLDDVGE	397
Db	430	RGISGEQRIHLLERIGIEGHOKTQGLRAIMNNRELLDDCOQVAVSGLICVALQLDDVGE	489
Qy	398	SVAPENQTLTGILAAVFVHQLTPRGVVRCLNLEBRVYLKPCRNAVGVNWRKSPDGD	457
Db	490	SVAPENQTLTGILAAVFVHQLTPRGVVRCLNLEBRVYLKPCRNAVGVNWRKSPDGD	549
Qy	458	DLMOGLGESEIRALFHHNNILLPSDHCEYYTFPHLSLQDFCAALYYVLGELIEIPALCP	517
Db	550	DLMOGLGESEIRALFHHNNILLPSDHCEYYTFPHLSLQDFCAALYYVLGELIEIPALCP	609

RESULT 8  
 US-10-677-943-24  
 Sequence 24, Application US/10677943  
 Publication No. US20040072297A1  
 GENERAL INFORMATION:  
 APPLICANT: The Government of the United States of America as  
 APPLICANT: represented by the Secretary of the Department of Health and  
 APPLICANT: Human Services  
 APPLICANT: Nelson, Lawrence  
 APPLICANT: Tong, Zhi-Bin  
 TITLE OF INVENTION: Human Gene Critical to Fertility  
 FILE REFERENCE: 4239-64790  
 CURRENT APPLICATION NUMBER: US/10/677,943  
 CURRENT FILING DATE: 2003-10-01  
 PRIOR APPLICATION NUMBER: 60/241,510  
 PRIOR FILING DATE: 2000-10-18  
 PRIOR APPLICATION NUMBER: PCT/US02/09776  
 PRIOR FILING DATE: 2002-03-29  
 PRIOR APPLICATION NUMBER: PCT/US01/10981  
 PRIOR FILING DATE: 2001-04-04  
 NUMBER OF SEQ ID NOS: 42  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO 24  
 LENGTH: 1200  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-677-943-24

QY	518	YVVEKTKRSMELKQAGPHIHSILMMKRFLFGVSESDVRPRLPLEYJLACPVJLGVKOKLIMV	577
Db	610	LYVEKTKRSMELKQAGPHIHSILMMKRFLFGVSESDVRPRLPLEYJLACPVJLGVKOKLIMV	669
QY	578	SLGQGPNAITPGDITLDAFHCLFETODEFVRLALNSFOEVLPIINONIDLIASSFCLOH	637
Db	670	SLGQGPNAITPGDITLDAFHCLFETODEFVRLALNSFOEVLPIINONIDLIASSFCLOH	729
QY	638	CPYLKINVDYKGIFFPDESABACPVVPLMMRDKTLISEQWEDFCSMLGTHPLKQLDLC	697
Db	730	CPYLKIRIVDYKGIFFPDESABACPVVPLMMRDKTLISEQWEDFCSMLGTHPLKQLDLC	789
QY	698	SSILITERAMKTLCAQLRHPTCKIOTLMMRNNOITPGVOHWRIVWANRNLRSNLNGTTL	757
Db	790	SSILITERAMKTLCAQLRHPTCKIOTLMMRNNOITPGVOHWRIVWANRNLRSNLNGTTL	849
QY	758	KEEDVRMACEALKHPKCLLESIRLDDCCGLTHACVYLKISQIILTPSPSLKSLIAGNKVYDQ	817
Db	850	KEEDVRMACEALKHPKCLLESIRLDDCCGLTHACVYLKISQIILTPSPSLKSLIAGNKVYDQ	909
QY	818	GMPPLSDALRYSOCLQYLIEDCGITATGCSLASALVSNRSLTHLCLSNNSLQNEGVN	877
Db	910	GVTPLSDALRYSOCLQYLIEDCGITATGCSLASALVSNRSLTHLCLSNNSLQNEGVN	969
QY	878	LLCSRMRLPHCSLORIMLNOCHLDPAGGFLALALMGNSWITLHLSLNNPVEDNGVKLLC	937
Db	970	LLCSRMRLPHCSLORIMLNOCHLDPAGGFLALALMGNSWITLHLSLNNPVEDNGVKLLC	1022
QY	938	EWMEPSCHLDLELVKCHLTPAACCESLSCVYSRSRHLKSLDLTDNALGDGVAALCEGL	997
Db	1030	EWMEPSCHLDLELVKCHLTPAACCESLSCVYSRSRHLKSLDLTDNALGDGVAALCEGL	1088
QY	998	KOKRSVTLRLGLKACGLTSDCCCEALSLASCNRHLTSLNLYONNNSPRGMMKLCAPACP	1057
Db	1090	KOKRSVTLRLGLKACGLTSDCCCEALSLASCNRHLTSLNLYONNNSPRGMMKLCAPACP	1144
QY	1058	TSNIOIIGLMMKQYVPOVIRKLLSEYQVLKPRVVIDGSMHSPDEDDR 1103	
Db	1150	TSNIOIIGLMMKQYVPOVIRKLLSEYQVLKPRVVIDGSMHSPDEDDR 1195	
<p>RESULT 9</p> <p>US-10-860-761-4</p> <p>; Sequence 4, Application US/10860761</p> <p>; Publication No. US2004024875A1</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: WYETH</p> <p>; TITLE OF INVENTION: METHODS FOR SCREENING INHIBITORS OF APOPTOSIS</p> <p>; FILE REFERENCE: AM101318</p> <p>; CURRENT APPLICATION NUMBER: US/10/860,761</p> <p>; CURRENT FILING DATE: 2004-06-03</p> <p>; NUMBER OF SEQ ID NOS: 31</p> <p>; SOFTWARE: PatentIn version 3.2</p> <p>; SEQ ID NO 4</p> <p>; LENGTH: 1200</p> <p>; TYPE: PRT</p> <p>; ORGANISM: Homo sapiens</p> <p>US-10-860-761-4</p>			
<p>Query Match 44.8%; Score 602; DB 16; Length 1200;</p> <p>Best Local Similarity 99.6%; Pred. No. 0;</p> <p>Matches 1002; Conservative 0; Mismatches 4; Indels 0; Gaps 0;</p>			
QY	98	ISQAMEQBGATAETBEQEIISQAMEQBGATAETBEQCHGDITMDYKSHVWTKPFAEBEDV	157
Db	190	ISQAMEQBGATAETBEQEIISQAMEQBGATAETBEQCHGDITMDYKSHVWTKPFAEBEDV	249
QY	158	RRSPENTAADPEMOTLAGAFDSDRWGPRPTVVLHGSGLGSKSALARIVLCMAQGGGLY	217
Db	250	RRSPENTAADPEMOTLAGAFDSDRWGPRPTVVLHGSGLGSKSALARIVLCMAQGGGLY	309
QY	218	QGMSSYFFLPEVRMOKKSSVTEFISREWPDSQAPVTEIMSRBERLLFTIIDGFDLGS	277

Db	790	SSILTERAMTKLCAALRHPTCKIQTLAMFRNAQITPGVQHLWRIWMANRRLRSNLNGTTL	849
QY	758	KEEDVYMACEALKEPKCLLESRLDCCGLTHACYIKISQILTTSPSLKSLSLAGNKVTDQ	817
Db	850	KEEDVYMACEALKEPKCLLESRLDCCGLTHACYIKISQILTTSPSLKSLSLAGNKVTDQ	909
QY	818	GWPLSDRLARYSQCLQKTLIEDGSIITATGOSLASALVSNRSLTHLCLSNNSLGNEGN	877
Db	910	GVTPLSDRLARYSQCLQKTLIEDGSIITATGOSLASALVSNRSLTHLCLSNNSLGNEGN	969
QY	878	LLCSRMRLPHCSLORLMLNQCHLDTAGCGFLALALMGNSWLTHLSLNNPVEDNGVKLLC	937
Db	970	LLCSRMRLPHCSLORLMLNQCHLDTAGCGSLALALMGNSWLTHLSLNNPVEDNGVKLLC	1022
QY	938	EWMEPSCHLQDLBILVKCHLTAAACESLSCVISRSRHLKSLDLTNALGDGGVAALCEGL	997
Db	1030	EWMEPSCHLQDLBILVKCHLTAAACESLSCVISRSRHLKSLDLTNALGDGGVAALCEGL	1088
QY	998	KOKNSVTLRLSLKACGLTSDCCCEALSLALSCNRHLTSLNLYONNPSPKGMKMLCSAFACP	105
Db	1090	KOKNSVTLRLSLKACGLTSDCCCEALSLALSCNRHLTSLNLYONNPSPKGMKMLCSAFACP	1144
QY	1058	TSNLIQIGLAKMOYPVQIRKLLEEVQILKPRVVLDGSMHSPDEDDR	1103
Db	1150	TSNLIQIGLAKMOYPVQIRKLLEEVQILKPRVVLDGSMHSPDEDDR	1195

```

Db 310 QGMEYVFELPVREMRKKSSEVTEFISREWPDSQAPVTEIMSRPERLFTIDGFDLGS 369
Qy 278 VLNNDTKLCKDMAEKOPPTLIRSLRKVLLPESFLIVTVADVGEKLSKSEVSPRYLLV 337
Db 370 VLNNDTKLCKDMAEKOPPTLIRSLRKVLLPESFLIVTVADVGEKLSKSEVSPRYLLV 429
Qy 338 RGISGEORIHLLERIGIEGHOKTQGLRAIMNNRELLDQCOVPAYGSLICVALQLODVVGE 397
Db 430 RGISGEORIHLLERIGIEGHOKTQGLRAIMNNRELLDQCOVPAYGSLICVALQLODVVGE 489
Qy 398 SVAPNQTLLTGLHAAFFHOLTFRGVVRCNLEERVVLKRCMAVAGVNNRKSVPDGD 457
Db 490 SVAPNQTLLTGLHAAFFHOLTFRGVVRCNLEERVVLKRCMAVAGVNNRKSVPDGD 549
Qy 458 DLMVQGLSESELRALFHMNIILPDSHCEYVTFPHLSLODFCALYYVLBGLTEIPALCP 517
Db 550 DLMVQGLSESELRALFHMNIILPDSHCEYVTFPHLSLODFCALYYVLBGLTEIPALCP 609
Qy 518 LYVEKTKRSMELKQAGFHIHSLMMKRFGLVSESVDRPLEVLGCPVPLGVKQKLLHMV 577
Db 610 LYVEKTKRSMELKQAGFHIHSLMMKRFGLVSESVDRPLEVLGCPVPLGVKQKLLHMV 669
Qy 578 SILGQOPNATTPGDTLDAFHCLFETQDKEFVRLALNSFOEYVWLPINQMLDIASSFCLOH 637
Db 670 SILGQOPNATTPGDTLDAFHCLFETQDKEFVRLALNSFOEYVWLPINQMLDIASSFCLOH 729
Qy 638 CPYLKIKIVDVYKGIIPRDESAACVPLMMKDKLIEBOWEDFCSMGTGPHRLQDLG 697
Db 730 CPYLKIKIVDVYKGIIPRDESAACVPLMMKDKLIEBOWEDFCSMGTGPHRLQDLG 789
Qy 698 SSILTERAMKTLCAKLRHPTCKIQTLMPRNAQITPGVOHLMRIWANNENLSLNGTHTL 757
Db 790 SSILTERAMKTLCAKLRHPTCKIQTLMPRNAQITPGVOHLMRIWANNENLSLNGTHTL 849
Qy 758 KEEVDYRMACEALKHPKCLIESRLDCCGITHACYIKISQIILTTSPSLKSLIAGNKVTDQ 817
Db 850 KEEVDYRMACEALKHPKCLIESRLDCCGITHACYIKISQIILTTSPSLKSLIAGNKVTDQ 909
Qy 818 GMPPLSDALRVSGCALOKLILIEDCGITATGCGSLASALVSNRSLTHLCLSNNSLGENEYV 877
Db 910 GMPPLSDALRVSGCALOKLILIEDCGITATGCGSLASALVSNRSLTHLCLSNNSLGENEYV 969
Qy 878 LLCRSMLPHCSLQRLMLNOCHLDTAGCGFLATLALMGNSWLTHTLSLNMNPEVDNGVKKL 937
Db 970 LLCRSMLPHCSLQRLMLNOCHLDTAGCGFLATLALMGNSWLTHTLSLNMNPEVDNGVKKL 1029
Qy 938 EYMRPESCHLODLVYKCHLTACCESLSCVTSRSHLKSJDLITNALGDGVAAALCEGL 997
Db 1030 EYMRPESCHLODLVYKCHLTACCESLSCVTSRSHLKSJDLITNALGDGVAAALCEGL 1089
Qy 998 KOKNSVLRRLGKAGLTSDCCEALSLALSCNRHLTSINLVONNSPKGMMLCSAFACP 1057
Db 1090 KOKNSVLRRLGKAGLTSDCCEALSLALSCNRHLTSINLVONNSPKGMMLCSAFACP 1149
Qy 1058 TSNLQIIGLMKQYEVQIRKLEEVQILKPRVVDIGSMHSPEDDR 1103
Db 1150 TSNLQIIGLMKQYEVQIRKLEEVQILKPRVVDIGSMHSPEDDR 1195

```

```

RESULT 10
US-10-407-866-96
; Sequence 96, Application US/10407866
; Publication No. US20040002593A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: PAD Domain-Containing Polypeptides,
; FILE REFERENCE: 66554-10 (I/J 5755)
; CURRENT APPLICATION NUMBER: US/10/407,866
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 60/370,538

```

```

; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 682
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-866-96

```

```

Query Match 40.5%; Score 544; DB 15; Length 682;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 98 ISQAMEGATTAETEOEISQAMEGATTAETEOGHGSDTWDYKSHWTKFAEBDV 157
Db 139 ISQAMEGATTAETEOEISQAMEGATTAETEOGHGSDTWDYKSHWTKFAEBDV 198
Qy 158 RRSFENTADMPENQTTAGAFSDRWGRRPRTVVLHGSGIGKALARRIYLCMAQGLY 217
Db 199 RRSFENTADMPENQTTAGAFSDRWGRRPRTVVLHGSGIGKALARRIYLCMAQGLY 258
Qy 218 QGMEYVFELPVREMRKKSSEVTEFISREWPDSQAPVTEIMSRPERLFTIDGFDLGS 277
Db 259 QGMEYVFELPVREMRKKSSEVTEFISREWPDSQAPVTEIMSRPERLFTIDGFDLGS 318
Qy 278 VLNNDTKLCKDMAEKOPPTLIRSLRKVLLPESFLIVTVADVGEKLSKSEVSPRYLLV 337
Db 319 VLNNDTKLCKDMAEKOPPTLIRSLRKVLLPESFLIVTVADVGEKLSKSEVSPRYLLV 378
Qy 338 RGISGEORIHLLERIGIEGHOKTQGLRAIMNNRELLDQCOVPAYGSLICVALQLODVVGE 397
Db 379 RGISGEORIHLLERIGIEGHOKTQGLRAIMNNRELLDQCOVPAYGSLICVALQLODVVGE 438
Qy 398 SVAPNQTLLTGLHAAFFHOLTFRGVVRCNLEERVVLKRCMAVAGVNNRKSVPDGD 457
Db 439 SVAPNQTLLTGLHAAFFHOLTFRGVVRCNLEERVVLKRCMAVAGVNNRKSVPDGD 498
Qy 458 DLMVQGLSESELRALFHMNIILPDSHCEYVTFPHLSLODFCALYYVLBGLTEIPALCP 517
Db 499 DLMVQGLSESELRALFHMNIILPDSHCEYVTFPHLSLODFCALYYVLBGLTEIPALCP 558
Qy 518 LYVEKTKRSMELKQAGFHIHSLMMKRFGLVSESVDRPLEVLGCPVPLGVKQKLLHMV 577
Db 559 LYVEKTKRSMELKQAGFHIHSLMMKRFGLVSESVDRPLEVLGCPVPLGVKQKLLHMV 618
Qy 578 SILGQOPNATTPGDTLDAFHCLFETQDKEFVRLALNSFOEYVWLPINQMLDIASSFCLOH 637
Db 619 SILGQOPNATTPGDTLDAFHCLFETQDKEFVRLALNSFOEYVWLPINQMLDIASSFCLOH 678
Qy 638 CPYL 641
Db 679 CPYL 682

```

```

RESULT 11
US-10-399-443-2
; Sequence 2, Application US/10399443
; Publication No. US2004002869A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as Represented by the
; APPLICANT: Secretary, Department of Health & Human Services, The National Institute
; APPLICANT: Health
; APPLICANT: Nelson, Lawrence M.
; APPLICANT: Tong, Zhi-Bin
; APPLICANT: Nelson, Lawrence
; APPLICANT: Zhi-Bin, Tong
; TITLE OF INVENTION: Human Gene Critical to Fertility
; FILE REFERENCE: 4239-64785
; CURRENT APPLICATION NUMBER: US/10/399,443
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/241,510
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/10981

```

;; PRIOR FILING DATE: 2001-04-04  
;; NUMBER OF SEQ ID NOS: 24  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 2  
;; LENGTH: 385  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-399-443-2

Query Match 11.8%; Score 158; DB 15; Length 385;  
Best Local Similarity 100.0%; Pred. No. 1.1e-142;  
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 AAKOPFTLIRSLRKVLLPESFLITVYADVGEKLSKSVSPRYLVVVGISEGRIHLL 349  
DB 39 AAKOPFTLIRSLRKVLLPESFLITVYADVGEKLSKSVSPRYLVVVGISEGRIHLL 98  
QY 350 LERGI GEHOKTQGLRAIMNNRELLDQCQVPAVGSLLICVALQIDVVGESVAPFNQTLTGL 409  
DB 99 LERGI GEHOKTQGLRAIMNNRELLDQCQVPAVGSLLICVALQIDVVGESVAPFNQTLTGL 158  
QY 410 HAAFFVHQLTPRGVVRRCINLEERVVLKRFCCMAVEGV 447  
DB 159 HAAFFVHQLTPRGVVRRCINLEERVVLKRFCCMAVEGV 196

RESULT 12  
US-10-677-943-2  
;; Sequence 2, Application US/10677943  
;; Publication No. US20040072297A1

;; GENERAL INFORMATION:  
;; APPLICANT: The Government of the United States of America as  
;; APPLICANT: represented by the Secretary of the Department of Health and  
;; APPLICANT: Human Services  
;; APPLICANT: Nelson, Lawrence  
;; APPLICANT: Tong, Zhi-Bin  
;; TITLE OF INVENTION: Human Gene Critical to Fertility  
;; FILE REFERENCE: 4239-64790  
;; CURRENT APPLICATION NUMBER: US/10/677,943  
;; CURRENT FILING DATE: 2003-10-01  
;; PRIOR APPLICATION NUMBER: 60/241,510  
;; PRIOR FILING DATE: 2000-10-18  
;; PRIOR APPLICATION NUMBER: PCT/US02/09776  
;; PRIOR FILING DATE: 2002-03-29  
;; PRIOR APPLICATION NUMBER: PCT/US01/10981  
;; PRIOR FILING DATE: 2001-04-04  
;; NUMBER OF SEQ ID NOS: 42  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 2  
;; LENGTH: 385  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-677-943-2

Query Match 11.8%; Score 158; DB 15; Length 385;  
Best Local Similarity 100.0%; Pred. No. 1.1e-142;  
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 AAKOPFTLIRSLRKVLLPESFLITVYADVGEKLSKSVSPRYLVVVGISEGRIHLL 349  
DB 39 AAKOPFTLIRSLRKVLLPESFLITVYADVGEKLSKSVSPRYLVVVGISEGRIHLL 98  
QY 350 LERGI GEHOKTQGLRAIMNNRELLDQCQVPAVGSLLICVALQIDVVGESVAPFNQTLTGL 409  
DB 99 LERGI GEHOKTQGLRAIMNNRELLDQCQVPAVGSLLICVALQIDVVGESVAPFNQTLTGL 158  
QY 410 HAAFFVHQLTPRGVVRRCINLEERVVLKRFCCMAVEGV 447  
DB 159 HAAFFVHQLTPRGVVRRCINLEERVVLKRFCCMAVEGV 196

RESULT 13  
US-10-399-443-4

;; Sequence 4, Application US/10399443  
;; Publication No. US20040028669A1

;; GENERAL INFORMATION:  
;; APPLICANT: The Government of the United States of America, as Represented by the  
;; APPLICANT: Secretary, Department of Health & Human Services, The National Institute  
;; APPLICANT: Health  
;; APPLICANT: Nelson, Lawrence M.  
;; APPLICANT: Tong, Zhi-Bin  
;; APPLICANT: Nelson, Lawrence  
;; APPLICANT: Zhi-Bin, Tong  
;; TITLE OF INVENTION: Human Gene Critical to Fertility  
;; FILE REFERENCE: 4239-64785  
;; CURRENT APPLICATION NUMBER: US/10/399,443  
;; CURRENT FILING DATE: 2003-04-16  
;; PRIOR APPLICATION NUMBER: 60/241,510  
;; PRIOR FILING DATE: 2000-10-18  
;; PRIOR APPLICATION NUMBER: PCT/US01/10981  
;; PRIOR FILING DATE: 2001-04-04  
;; NUMBER OF SEQ ID NOS: 24  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 4  
;; LENGTH: 258  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-399-443-4

Query Match 11.1%; Score 149; DB 15; Length 258;  
Best Local Similarity 100.0%; Pred. No. 3.6e-134;  
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 918 LTHLSLNNPVEDNKKLCEVMPRESCHDLELVKCHLTPACCESSJCVISRSRHKS 977  
DB 68 LTHLSLNNPVEDNKKLCEVMPRESCHDLELVKCHLTPACCESSJCVISRSRHKS 127  
QY 978 LDTLNNALGDGVVALCEGLKQKNSVLTRLGKACGLTSDCEALSLALSCNRHLSINTL 1037  
DB 128 LDTLNNALGDGVVALCEGLKQKNSVLTRLGKACGLTSDCEALSLALSCNRHLSINTL 187  
QY 1038 VQNNFSPKGMKLCGAFACPTSNLQITGL 1066  
DB 188 VQNNFSPKGMKLCGAFACPTSNLQITGL 216

RESULT 14  
US-10-677-943-4  
;; Sequence 4, Application US/10677943  
;; Publication No. US20040072297A1

;; GENERAL INFORMATION:  
;; APPLICANT: The Government of the United States of America as  
;; APPLICANT: represented by the Secretary of the Department of Health and  
;; APPLICANT: Human Services  
;; APPLICANT: Nelson, Lawrence  
;; APPLICANT: Tong, Zhi-Bin  
;; TITLE OF INVENTION: Human Gene Critical to Fertility  
;; FILE REFERENCE: 4239-64790  
;; CURRENT APPLICATION NUMBER: US/10/677,943  
;; CURRENT FILING DATE: 2003-10-01  
;; PRIOR APPLICATION NUMBER: 60/241,510  
;; PRIOR FILING DATE: 2000-10-18  
;; PRIOR APPLICATION NUMBER: PCT/US02/09776  
;; PRIOR FILING DATE: 2002-03-29  
;; PRIOR APPLICATION NUMBER: PCT/US01/10981  
;; PRIOR FILING DATE: 2001-04-04  
;; NUMBER OF SEQ ID NOS: 42  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 4  
;; LENGTH: 258  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-677-943-4

Query Match 11.1%; Score 149; DB 15; Length 258;  
Best Local Similarity 100.0%; Pred. No. 3.6e-134;

Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 918 LTHLSLNNPVEDNGVKLLCEVMRRPSCHLODLIELVKCHLTRACCESLSCVTSRSHLKS 977

DB 68 LTHLSLNNPVEDNGVKLLCEVMRRPSCHLODLIELVKCHLTRACCESLSCVTSRSHLKS 127

QY 978 LDLTNNALGDGGVALCEGLKOKNSVLTRLGLKACGLTSDCCCEALSLALSCNRHLSLNL 1037

DB 128 LDLTNNALGDGGVALCEGLKOKNSVLTRLGLKACGLTSDCCCEALSLALSCNRHLSLNL 187

QY 1038 VQNNFSPKGMKLCSAFACPTSNLQIIGL 1066

DB 188 VQNNFSPKGMKLCSAFACPTSNLQIIGL 216

# RESULT 15

US-09-965-621-63

; Sequence 63, Application US/09965621

; Publication No. US20030077699A1

; GENERAL INFORMATION:

; APPLICANT: Reed, John C.

; APPLICANT: Godzik, Adam

; APPLICANT: Chu, Zhi-Liang

; APPLICANT: Pawlowski, Krzysztof

; APPLICANT: Fiorentino, Ioredana

; APPLICANT: Ariza, Maria Eugenia

; APPLICANT: Stehlik, Christian

; TITLE OF INVENTION: PAD Domain-Containing Polypeptides,

; FILE REFERENCE: P-LJ 4816

; CURRENT APPLICATION NUMBER: US/09/965,621

; PRIOR FILING DATE: 2001-09-25

; PRIOR FILING DATE: 2000-09-26

; NUMBER OF SEQ ID NOS: 64

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 63

; LENGTH: 190

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-965-621-63

Query Match 0.9%; Score 12; DB 10; Length 190;

Best Local Similarity 100.0%; Pred. No. 0.019;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 PERLFTIDGFD 273

DB 77 PERLFTIDGFD 88

Search completed: July 19, 2005, 12:27:33  
Job time : 127 secs